

# Iconicity in Spatial Language Guides Visual Attention: A Comparison Between Signers' and Speakers' Eye Gaze During Message Preparation

Francie Manhardt  
Radboud University

Aslı Özyürek  
Radboud University and Max Planck Institute for  
Psycholinguistics, Nijmegen, the Netherlands

Beyza Sümer  
Radboud University and University of Amsterdam

Kimberley Mulder, Dilay Z. Karadöller,  
and Susanne Brouwer  
Radboud University

To talk about space, spoken languages rely on arbitrary and categorical forms (e.g., left, right). In sign languages, however, the visual–spatial modality allows for iconic encodings (motivated form–meaning mappings) of space in which form and location of the hands bear resemblance to the objects and spatial relations depicted. We assessed whether the iconic encodings in sign languages guide visual attention to spatial relations differently than spatial encodings in spoken languages during message preparation at the sentence level. Using a visual world production eye-tracking paradigm, we compared 20 deaf native signers of Sign-Language-of-the-Netherlands and 20 Dutch speakers' visual attention to describe left versus right configurations of objects (e.g., “pen is to the left/right of cup”). Participants viewed 4-picture displays in which each picture contained the same 2 objects but in different spatial relations (lateral [left/right], sagittal [front/behind], topological [in/on]) to each other. They described the target picture (left/right) highlighted by an arrow. During message preparation, signers, but not speakers, experienced increasing eye-gaze competition from other spatial configurations. This effect was absent during picture viewing prior to message preparation of relational encoding. Moreover, signers' visual attention to lateral and/or sagittal relations was predicted by the type of iconicity (i.e., object and space resemblance vs. space resemblance only) in their spatial descriptions. Findings are discussed in relation to how “thinking for speaking” differs from “thinking for signing” and how iconicity can mediate the link between language and human experience and guides signers' but not speakers' attention to visual aspects of the world.

*Keywords:* eye-tracking, iconicity, language production, sign language, visual attention

*Supplemental materials:* <http://dx.doi.org/10.1037/xlm0000843.supp>

As humans, we constantly explore the visual world around us with our eyes by selecting and attending to relevant details, while ignoring

the irrelevant ones. Over the past decades researchers have studied and claimed a link between eye-gaze patterns and production and comprehension of spoken language (for language comprehension, see e.g., Cooper, 1974; Tanenhaus & Trueswell, 2006; for language production within languages at the lexical level as well as for complex messages, see e.g., Griffin & Bock, 2000; Konopka & Meyer, 2014; van de Velde, Meyer, & Konopka, 2014; for language production across different languages, see e.g., Flecken, Von Stutterheim, & Carroll, 2014; Papafragou, Hulbert, & Trueswell, 2008; Slobin, 2003). This link has been extensively studied for spoken languages but we know little about whether these effects are similar between signed versus spoken languages or whether the modality (visual vs. acoustic) of language guides visual attention differently. Recent studies have provided evidence that eye-gaze and sign language comprehension are also linked as found in spoken languages (e.g., Lieberman, Borovsky, Hatrak, & Mayberry, 2015; Lieberman, Borovsky, & Mayberry, 2018; Thompson, Vinson, Fox, & Vigliocco, 2013). However, this evidence is limited to comprehension and it is not known whether the modality of expressions guides visual attention differently for linguistic production at the sentence level and during preparation of messages.

This article was published Online First April 30, 2020.

 Francie Manhardt, Centre for Language Studies, Radboud University; Aslı Özyürek, Centre for Language Studies and Donders Institute for Brain, Cognition and Behavior, Radboud University, and Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands; Beyza Sümer, Centre for Language Studies, Radboud University, and Department of Linguistics, University of Amsterdam; Kimberley Mulder, Dilay Z. Karadöller, and Susanne Brouwer, Centre for Language Studies, Radboud University.

This work has been supported by an NWO VICI grant awarded to Aslı Özyürek. We thank the Max Planck Institute for Psycholinguistics and our deaf and hearing assistants Tom Uittenbogert and Renske Schilte for their help collecting, annotating, and coding data, as well as our deaf and hearing addressees Marcel Koopmans and Julia Merkus. The authors are very grateful to Nick Wood and Jeroen Geerts for helping us with processing the video data.

Correspondence concerning this article should be addressed to Francie Manhardt, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6252 XD, Nijmegen, the Netherlands. E-mail: [francie.manhardt@mpi.nl](mailto:francie.manhardt@mpi.nl)

The present study investigates for the first time the link between eye-gaze and linguistic expressions in signers and speakers during preparation of messages at the sentence level. More specifically, it tests whether a modality-specific aspect of sign language encoding, that is iconicity in the expressions, guides visual attention differently for signers than for speakers during message preparation for spatial language (e.g., “the pen is to the left of the cup”). Sign languages, unlike spoken languages, encode spatial relations between entities in iconic ways where linguistic forms resemble the forms of the objects and the spatial relations between them. In the current study, speakers and signers are asked to view pictures depicting different spatial relations between items and describe one of them while their eye gaze patterns are recorded prior to and during message preparation. Eye movements are analyzed to see whether signers and speakers show similar or different eye-gaze patterns, and more specifically, if differences in visual attention can be linked to the iconic versus arbitrary and categorical ways of encoding information in sign versus spoken language.

### The Link Between Language and Eye Gaze in Spoken Languages

Previous research has shown that during language production speakers direct their eye gaze at the referents they are describing in the order that they mention them, reflecting the incremental characteristics of spoken language planning (Griffin, 2004; Griffin & Bock, 2000; Konopka & Meyer, 2014; Meyer, Sleiderink, & Levelt, 1998; van de Velde et al., 2014). Based on this evidence, it has been claimed that there is a tight link between incremental speech production and eye gaze (i.e., speech–gaze link). Furthermore, linguistic variation across different languages in which different elements of a scene are encoded seems to guide speakers’ visual attention to different components of these visual scenes during message preparation (i.e., thinking for speaking, see Bunker, Skordos, Trueswell, & Papafragou, 2016; Flecken, Carroll, Weimar, & Von Stutterheim, 2015; Flecken et al., 2014; Papafragou et al., 2008; Slobin, 2003; Trueswell & Papafragou, 2010). For example, Papafragou and colleagues (2008) used eye tracking to show that language influences visual attention during message preparation. The study assessed whether cross-linguistic differences between English and Greek, in how a motion event is encoded, directs visual attention to different parts of the event. Greek predominantly encodes the Path (i.e., the direction) of the motion in the main verb (Talmy, 1985; Talmy, 2003), whereas English typically encodes Manner of the motion (i.e., how the motion is performed) in the main verb. During planning of event descriptions (e.g., someone is running up the stairs), Greek speakers encoded and paid more attention to the Path (e.g., ascending the stairs), while English speakers encoded and paid more attention to the Manner of the event (e.g., running). Therefore, eye-gaze patterns differed between Greek and English speakers during message preparation. Similarly, categorization differences between German and Korean speakers impacted the way these speakers visually attended to spatial relations between two objects (Goller, Lee, Ansoerge, & Choi, 2017). Korean speakers based their linguistic categorization of space on the degree of fit between the two objects (e.g., *kkita*

for loose fit vs. *netha* for tight fit) whereas German speakers did not. Instead, German differentiates between support and containment (e.g., *auf* for support vs. *in* for containment). Eye-gaze patterns indicated that Korean speakers looked equally likely at the two objects, while German speakers looked more frequently at the ground object. Overall, previous research suggests that during the planning of describing spatial scenes, speakers of different spoken languages focus on those aspects of scenes that are relevant for linguistic encoding. However, this body of evidence for language production influencing visual attention is limited to the study of spoken languages. It yet needs to be investigated whether “thinking for speaking” (Slobin, 2003) might differ from “thinking for signing” because of the iconic ways sign languages are organized for expressing spatial relations, unlike spoken languages.

### Iconicity in Sign Languages and Effects in Language Processing

Some aspects of sign languages differ from spoken languages in terms of iconicity which is defined as the motivated *mapping* between meaning and a visual or auditory linguistic form (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Emmorey, 2014; Perniss, Thompson, & Vigliocco, 2010). In spoken languages, arbitrariness mostly governs the relation between form and meaning (but see Assaneo, Nichols, & Trevisan, 2011; Dingemanse, 2012 for an overview of existing iconic [i.e., motivated] forms in various spoken languages). In contrast, the visual nature of the modality of sign languages allows a large proportion of motivated one-to-one mappings in their linguistic structures. Within sign languages, iconicity can be found at the lexical level (e.g., in the sign for airplane in Sign Language of the Netherlands (*Nederlandse Gebarentaal*, NGT) the extended thumb and pinkie depict the wings and the movement of the sign depicts the movement up into the sky) and also at the sentence level in which signers can map visual features of spatial relations between entities onto their hands and signing space.

While the investigation on the influence of iconicity on spatial language processing in sign languages has just begun, a number of behavioral studies suggest that iconicity at the lexical level can affect sign language processing during language comprehension (Grote & Linz, 2003; Thompson, Vinson, & Vigliocco, 2009, 2010; Thompson, Vinson, Woll, & Vigliocco, 2012). For instance, Grote and Linz (2003) found that native signers of German Sign Language (*Deutsche Gebärdensprache*, DGS) were faster in judging picture–sign relationships for iconic properties of signs. That is, signers were faster to judge object–property relations when the iconic property of the sign (e.g., tracing the beak of the eagle for the DGS sign for eagle) matched the following picture compared to when it did not match (e.g., a picture of a beak vs. a picture of a wing). Interestingly, this effect was not found in speakers of German, who received a version of the same task containing the written translation equivalents in German. In another study, signers of British Sign Language (BSL) were slower in deciding whether a sign involved a straight or bent handshape (i.e., phonological parameters of a sign) when signs were iconic compared to when they were arbitrary (Thompson et al., 2010). Moreover, it has been claimed that iconicity is not only relevant for language

processing but can actually mediate the link between language and human experience (i.e., action and perception) because sign languages' visuospatial mappings of many signs are salient aspects of the mental representations of their corresponding entities in the world (Perniss & Vigliocco, 2014; Thompson et al., 2012; Vigliocco, Perniss, & Vinson, 2014; Vigliocco, Vinson, Woolfe, Dye, & Woll, 2005; Vinson, Thompson, Skinner, & Vigliocco, 2015). However, the role of iconicity in semantic language processing at the lexical level is actively debated since many studies have also failed to show iconicity effects on language comprehension (Bellugi & Klima, 1976; Bosworth & Emmorey, 2010). These findings suggest that iconicity effects on language comprehension depend on the investigated sample (e.g., native vs. late signers), the task, and the type of iconicity. For sign language production, however, Navarrete, Peressotti, Lerose, and Miozzo (2017) presented evidence for a facilitating effect of iconicity during a target-distractor picture naming task. That is, picture distractors with iconic signs induced faster responses than when picture distractors were arbitrary, suggesting that iconicity can drive the activation of unproduced lexical signs and can therefore affect linguistic production planning.

Overall, previous studies have shown that iconicity can affect sign language processing during language comprehension and production, although the latter has been less studied. Furthermore, these findings are limited to the lexical level. To our knowledge, there are no studies so far looking at whether and how message preparation at the sentence level guides visual attention. Within the domain of spatial language (e.g., the pen is to the left of the cup) and sign languages, relations between two entities and their similarity to the linguistic forms are important (i.e., iconicity). This allows us to assess how visual attention relates to visual similarity between linguistic forms and visual relations between two objects and shape features of the objects. Therefore, the present study aims to investigate whether and how the encoding of iconic forms in sign languages at the sentence level guides visual attention to spatial relations differently than spoken language production during message preparation.

### The Present Study

As a consequence of the vocal modality, speech forms in spoken languages rely on encoding three-dimensional events and spatial relations into arbitrary and categorical forms (see also Perniss, Zwitserlood, & Özyürek, 2015). The visual-spatial nature of the modality of sign languages, however, allows iconic representations of space by directly mapping entities and their spatial relations onto the hands and the signing space (Emmorey, 1996, 2002; Emmorey, McCullough, Mehta, Ponto, & Grabowski, 2013; Perniss, Özyürek, & Morgan, 2015; Talmy, 2003). In the present study, we investigated whether these iconic spatial encodings in sign languages guide signers' attention to spatial relations differently than that of speakers during message preparation. That is, whether thinking for signing, thus the way iconic forms depict aspects of referents and relations, differs from thinking for speaking (Slobin, 2003). In addition, we aim to explore whether the type of iconicity within signers' spatial encodings influences their visual attention differently. Below, we first give some more detailed information about iconicity and different types of iconicity in sign languages to describe spatial relations.

### Iconicity in Encoding Spatial Relations in Sign Languages

Sign languages use iconic forms to encode spatial relations unlike spoken languages. In particular, different constructions of spatial languages in sign languages exploit different types of iconic encodings. The common constructions to express spatial relations between entities are called *classifier constructions* (CLs). In using CLs, signers predominantly map shape and orientation properties of objects and relations among them onto their hands and the signing space in front of them by placing both hands in front of the body (e.g., Emmorey, 1996, 2002; Perniss, Özyürek et al., 2015; Zwitserlood, 2012). For example, in NGT signers may first introduce the lexical signs of the two entities (e.g., the pen and the cup), followed by a round handshape to represent the shape of the cup and the index finger to represent the thin, elongated shape of a pen, placing both hands next to each other to match the signers' view of the relative relations of the entities to each other (see Figure 1A). These mappings "mirror" or visually resemble both the object properties (i.e., shape) as well as the spatial relations between entities onto the signing space as viewed from the signers' own perspective (Emmorey & Herzig, 2003; Perniss et al., 2015; Perniss et al., 2010).

However, there are other iconic forms that signers can use, which encode only the spatial relation without encoding specific information about object properties (e.g., shape; *relational lexemes* [RLs]; e.g., Arik, 2013; Perniss, Zwitserlood, et al., 2015; Sümer,

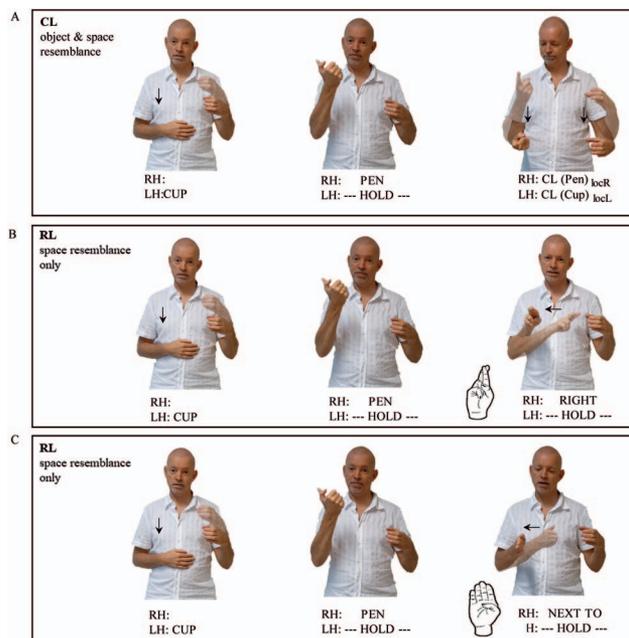


Figure 1. Types of iconicity in different linguistic expressions for "pen is to the right of the cup" in NGT. Panel A illustrates a classifier construction (CL) that visually resembles both object properties and space. Panel B and C demonstrate relation lexemes (RL) for right (R-handshape; B) and next to (to the right; B-handshape; C), which resemble space only. RH stands for right hand, LH for left hand, and loc for locative placement (to the right or left). These images are used with permission. See the online article for the color version of this figure.

Permiss, Zwitserlood, & Özyürek, 2014). RLs thus differ from CLs in their type of iconicity. RLs resemble space (i.e., figure object to the right or left of the ground object), while CLs resemble object properties and space (for neurobiological evidence for different processing of CLs vs. RLs in signers see Emmorey et al., 2013). Specifically, for left/right spatial configurations in NGT, there are two forms of RLs available to encode spatial relations on the lateral axis. First, NGT signers can use RLs for left and right (Figure 1B). These lexemes contain first letters of the spoken words (i.e., *L* handshape for Dutch left: *links*; *R* handshape for Dutch right: *rechts*) as well as the hand movement to the right or to the left (corresponding to the position of the pen in regard to the cup). Second, NGT signers can use a RL meaning *next to* (Figure 1C) consisting of a *B*-handshape moving to the right or left (corresponding to the location of the figure object in regard to the ground object, but without changing the *B*-handshape; see Permiss, 2007; Permiss et al., 2015 for a similar relational lexeme in DGS and Sümer, 2015, in Turkish Sign Language, *Türk İşareli Dili*; TİD). Both forms of RLs in NGT express the relative spatial relation between two entities in iconic ways (one-to-one mappings of the pictures) through space resemblance but not through resemblance to object properties, unlike CLs. Therefore, the types of iconicity in spatial language encoding can differ and raise novel questions whether these different types of iconic forms selected guide visual attention to spatial relations differently.

### Visual World Production Eye-Tracking Paradigm

In order to assess whether the above-mentioned types of iconic forms in NGT influence eye gaze to relations between two objects differently than spoken Dutch during message preparation, we used a visual world eye-tracking paradigm. Instead of assessing language comprehension, we used this paradigm for language production in a novel way (for a similar approach, see Davies & Kreysa, 2017, 2018). We presented speakers and signers with the traditional four-picture displays, but as different from previous research, each picture contained the same two objects but in different spatial relations to each other (i.e., left, right, front, behind, in, or on) as we were interested in relational encoding. After an initial viewing phase of the four pictures, participants saw a visual cue in the form of an arrow indicating the target picture. This visual cue then disappeared and the four pictures remained on the screen until a gray screen appeared signaling that it was time to describe (i.e., speaking/signing) the target picture. We recorded speakers' and signers' eye movements during two crucial display phases. First, we assessed the initial viewing phase before the target was indicated (i.e., *prearrow window*, Figure 2) serving as a baseline to explore whether speakers and signers observe the visual displays differently prior to message preparation for encoding of spatial relations (e.g., due to deafness, enhanced peripheral vision or sign language experience) and also to check whether our displays had intrinsic features that guided visual attention. Second, we assessed the viewing phase after the arrow disappeared until the production screen appeared (i.e., *postarrow window*, Figure 2), which was our main interest to investigate visual attention during message preparation.

In the four-picture displays we manipulated the ways in which spatial relations in nontarget pictures (i.e., competitors and distractors) competed with the target picture (i.e., right/left) either due

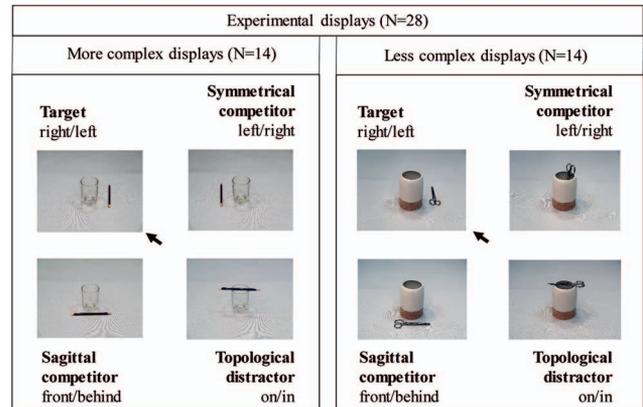


Figure 2. Examples of experimental displays (specified for more complex and less complex displays). See the online article for the color version of this figure.

to being *visually similar* and/or *semantically relevant* to the target picture (see Figure 2). In this manipulation, when nontarget pictures depicted figure objects placed to the sides of the ground object (i.e., left/right/front/behind; lateral and sagittal relations) they were considered more visually similar to the target than when they were in in/on relations (i.e., topological relations). The reason to consider lateral and sagittal relations as more similar to the target compared to topological relations is based on the view that the former is viewpoint dependent and therefore require perspective taking from the language user (Levinson, 1996, 2003). Thus, sagittal and lateral relations share more similarities to each other and to the target object in terms of viewpoint dependency and perspective taking and differ in those terms from topological relations, which are not viewpoint dependent and therefore do not require perspective taking. In addition to visual similarity, nontarget pictures were also defined as semantically relevant if they shared the same lateral symmetrical axis as the target object (i.e., left/right are more semantically related than left/front or left/in).

Following this approach, displays included (a) a *symmetrical competitor* (i.e., left/right), which was visually similar to the target picture and also semantically relevant, (b) a *sagittal competitor* (i.e., front/behind), which was only visually similar to the target picture (i.e., figure object is on the sagittal side of the ground object) but not semantically relevant, and (c) a *topological distractor* (i.e., in/on), which was neither visually nor semantically relevant. We reasoned that if iconicity of relational encodings matters, then the competitors would attract signers' attention more than speakers' and also where signers allocate their visual attention would depend on the type of iconicity used in their spatial encodings.

We further manipulated the complexity of displays in two ways: we either included the symmetrical competitor, creating more complex displays, or excluded it, creating less complex displays (for a similar approach using a contrast manipulation, see Davies & Kreysa, 2017, 2018; Sedivy, Tanenhaus, Chambers, & Carlson, 1999). We therefore decreased display complexity when a symmetrical competitor was absent due to its high degree of visual similarity to the left/right target picture. Thus, we assumed that the presence of a symmetrical competitor is more perceptually complex (i.e., more complex displays) compared to when it is absent

(i.e., less complex displays). In these so-called less complex displays we presented the target, one sagittal competitor, and two topological distractors. This manipulation served as a baseline for exploring whether speakers' and signers' linguistic descriptions and eye gaze are further influenced by the complexity of displays.

### Predictions

For linguistic productions, we predicted that speakers will use the most informative strategy available, independent of the complexity of displays (Davies & Kreysa, 2017). That is, speakers will prefer to use specific spatial relation nouns (e.g., left/right) over general spatial relation nouns (e.g., next to) despite the presence or absence of a symmetrical competitor. For signers, we predicted a preference of using CLs over RLs as demonstrated previously for American Sign Language (ASL, Emmorey, 2002), DGS (Perniss, 2007), and TİD (Sümer, 2015). It is less known how these strategies might differ in different complexity contexts. Due to the high level of informativeness of CLs (resembling objects and space) we predicted that signers prefer the most informative strategy (i.e., CLs) independent of the presence or absence of a symmetrical competitor.

For eye gaze prior to message preparation (prearrow window), we predicted that speakers and signers do not differ in the way they observe the visual displays. Previous research suggests that only during the planning of describing spatial scenes, speakers of different spoken languages look at those aspects of scenes differently and in ways that are relevant for linguistic encoding. During only viewing—that is, without the goal of language production—these cross-linguistic effects were absent (e.g., Papafragou et al., 2008; Slobin, 2003). In our study during the prearrow window, subjects do not know which spatial relation to encode. We therefore consider that eye-gaze patterns in the prearrow window do not reflect differences in encoding for spatial relations (i.e., the focus of the current study). We cannot rule out that eye gaze might be guided by preparation for object labels, similar in all of the four pictures, or other general visual attention differences between speakers and signers. However, and importantly, we did not expect differences between signers and speakers to arise prior to message preparation for neither more complex displays nor less complex display that would reflect differences in visual attention related to relational encodings.

For eye gaze during message preparation (postarrow window), we expected to find differences between speakers and signers related to relational encodings. More specifically, we predicted that in order to plan describing the relation depicted in the target picture signers experience more eye-gaze competition from the symmetrical competitor or the sagittal competitor than speakers. Moreover, this enhanced competition from signers might increase as time unfolds and gets close to the actual message encoding (unlike in prearrow window). Due to the visual similarity of those competitors, signers need to identify the relative locations and/or the orientation, size and shape of the objects in relation to each other to map information onto the signing space iconically (e.g., orientation/shape of figure hand placed in relation to hand resembling the ground object). Consequently, disambiguating where to move or place the hands in space might require more effort for signers during production planning, resulting in an increase in visual attention to the symmetrical and/or sagittal competitors in

order to describe the target picture. In contrast, for speakers, this kind of iconic information is not relevant for their categorical and arbitrary spatial expressions. Thus, we expected that speakers experience less competition than signers from the visually similar competitors (i.e., symmetrical and sagittal competitor). For the topological distractor, we did not expect differences in eye-gaze competition for both speakers and signers, since it is neither semantically relevant nor visually similar to the target picture. Following the same line of reasoning, for less complex displays we predicted that signers experience no enhanced eye-gaze competition from the topological competitor compared to speakers. However, we did predict more eye-gaze competition from the sagittal competitor for signers compared to speakers. We did, however, take into account the possibility that competitors might be viewed differently depending on the complexity of the displays.

Additionally, we also expected that within signers, the type of iconicity (CLs or RLs) preferred in their linguistic encodings would modulate visual attention differently, providing more evidence for linking gaze-allocation of signers to the iconicity of their encodings. As explained above, signers can either use CLs, which resemble both space and object properties in a one-to-one mapping (Figure 1A), or they could use RLs, in which iconicity focusses on spatial properties only (see Figure 1B, 1C). Depending on which strategy signers are planning to use we expected different eye-gaze patterns to arise. On the one hand, preparing to use CLs might lead to more eye-gaze competition from visually relevant pictures (i.e., symmetrical and/or sagittal competitor) because they require more effort to resemble the object properties and where to place the hands in space (e.g., placing a long-elongated handshape representing the pen to the left vs. right vs. front vs. behind of the round handshape representing the cup). On the other hand, planning to use RLs might lead to more competition from only the semantically relevant symmetrical competitor but not from the other visually similar sagittal competitor. RLs' iconicity focuses on the relative location in space in a more abstract and categorical fashion than CLs' and thus only semantically relevant competitors (i.e., symmetrical competitor) might elicit competition with the target picture. As above, we expected these types of competitions to change over time and increase getting close to the message encoding. We do not expect more eye-gaze competition from the topological distractor as it is neither visually nor semantically relevant for describing the left/right target configuration.

### Method

The method reported in this experiment was approved by the Humanities Ethics Assessment Committee of the Radboud University.

### Participants

The participants in this study were 20 native speakers of Dutch (11 female), and 20 deaf signers of NGT (16 females; see Table 1 for participants' descriptive statistics). To assess similarities in age and language proficiency across signers and speakers, we conducted Bayesian *t* tests in which we assessed the probability of the mean difference greater than zero and less than zero, using the R package *BayesianFirstAid* (Version 0.1; Bååth, 2014). The groups were similar in age (Bayesian two sample *t* test:  $M_{DIFF}(-5) > 0$ :

Table 1  
Descriptive Statistics for Signers and Speakers

Fixed effects	Speakers	Signers
Age (years)	32 (11)	37 (12)
Self-rated proficiency (comprehension)	4.87 (0.4)	4.80 (0.4)
Self-rated proficiency (production)	4.8 (0.4)	4.6 (0.8)
Visuo-spatial memory span	6.65 (1.1)	6.40 (1.1)

*Note.* Self-rated proficiency scores contain evaluations of Dutch from speakers and of *Nederlandse Gebarentaal* (NGT) from signers from 0 (*no knowledge*) to 5 (*native like*). Visuo-spatial memory span contains the average span per group (out of 9) assessed by the Corsi Block Tapping Task. *SDs* are reported in parentheses.

$p = 0.129$ ,  $M_{DIFF} (5) < 0$ :  $p = 0.871$ ). All signers were born to deaf parents and acquired NGT from birth (i.e., from their deaf signing parents), except one who acquired sign supported Dutch (i.e., manually coded form of Dutch, thus spoken Dutch is being supported by NGT signs) from birth. Two additional signers and one speaker were tested, but were excluded from the study due to another native language than NGT ( $N = 1$ ), or high errors in describing the pictures (signers:  $N = 1$ ; speakers:  $N = 1$ ). Additionally, one signer was discarded due to high eye-tracking loss (larger than 35%).

The self-rated language proficiency scores ranging from 0 (*no knowledge*) to 5 (*native like*) were collected from speakers for Dutch and signers for NGT separately for comprehension and production (for validity of self-ratings to measure proficiency, see Wilson, 1999). Comprehension scores of Dutch included scores for reading and listening, while the scores for NGT included understanding. Production scores of Dutch included speaking and writing, while the scores for NGT included signing. Bayesian two-sample  $t$  tests indicated similar comprehension scores ( $M_{DIFF} (-0.07) > 0$ :  $p = 0.501$ ,  $M_{DIFF} (0.07) < 0$ :  $p = 0.499$ ) and production scores ( $M_{DIFF} (-0.2) > 0$ :  $p = 0.502$ ,  $M_{DIFF} (0.2) < 0$ :  $p = 0.498$ ) between speakers and signers. Paired Bayesian  $t$  tests showed no difference between self-rated language production scores and comprehension scores within speakers ( $M_{DIFF} (-0.2) > 0$ :  $p = 0.506$ ,  $M_{DIFF} (0.2) < 0$ :  $p = 0.494$ ) and within signers ( $M_{DIFF} (-0.07) > 0$ :  $p = 0.503$ ,  $M_{DIFF} (0.07) < 0$ :  $p = 0.497$ ). We used a Corsi Block Tapping task (Corsi, 1972) to control for differences in visuospatial memory between signers and speakers, which yielded similarities between speakers' and signers' working memory span [Bayesian two sample  $t$  test:  $M_{DIFF} (-0.25) > 0$ :  $p = 0.726$ ,  $M_{DIFF} (0.25) < 0$ :  $p = 0.274$ ]. All Dutch speakers and the majority of the signers were tested at the Max Planck Institute for Psycholinguistics in Nijmegen, The Netherlands. Three signers were tested at their homes. Participation was voluntary and all participants received a financial compensation.

## Materials

Participants were presented with 84 visual displays containing four pictures (see Figure 2). Each picture displayed the same two objects in a spatial relation to each other (i.e., left, right, front, behind, on, in). A visual cue in the form of an arrow appeared in the middle of the screen and pointed at one of the four pictures, indicating the target picture. Twenty-eight visual displays were used as experimental trials, half of them containing more complex

displays and the other half containing less complex visual displays. The target in both conditions was always a left or right configuration. More complex displays included a symmetrical competitor (left or right, depending on whether the target was left or right), a sagittal competitor (front or behind) and a topological distractor (in or on). Less complex displays, importantly, did not contain a symmetrical competitor but included instead a sagittal competitor (front or behind) and two topological distractors (i.e., both in and on).

Next to the 28 experimental displays, we included 56 additional displays as filler trials to distract participants from the critical left/right configurations during the experiment. The filler displays consisted of four different patterns to offer visual variation and avoid biases to the left/right symmetry. Importantly, all filler displays contained front/behind/in/on targets instead of left/right, which were equally distributed across all 56 displays (i.e., targets:  $14 \times$  front,  $14 \times$  behind,  $14 \times$  in,  $14 \times$  on). Thus, in filler displays the arrow pointed at any other spatial relation but left/right. First, 14 filler displays contained the same combinations of spatial relations as the 14 experimental more complex displays; thus, they contained left, right, front/behind, in/on. Second, 14 filler displays resembled the 14 experimental less complex displays (i.e., containing left/right, front/behind, in, on). Third, 14 visual displays contained both a front and behind configuration to distract participants from our critical lateral contrast (i.e., front vs. behind, left/right, in/on). Finally, 14 visual displays also contained both front and behind relations, however, did not include any left/right spatial relations (i.e., front vs. behind, in, on).

Figure objects (e.g., the pen) were presented once, while Ground objects (e.g., the jar) occurred four times, but always in combination with another figure (e.g., jar–lemon, jar–watch, jar–screwdriver, and jar–jar lid). One Ground object was never presented more than twice in a row. Ground objects in experimental displays were always nonintrinsic objects to allow left/right interpretations rather than intrinsic interpretations. Additionally, target spatial relations (i.e., left/right) were presented not more than twice in a row to avoid biases to one type of spatial relations. The location of each configuration was counterbalanced across participants and trials. All visual displays have been piloted to assure that participants can name the presented objects.

## Procedure

Before the actual experiment, participants performed a familiarization task using displays similar to those in the actual eye-tracking experiment to introduce participants to the general complexity of our displays (i.e.,  $2 \times 2$  grid with two objects in different spatial relations to each other). After answering some questions about the displays, we proceeded with the actual eye-tracking description task. The experiment was preceded by three practice trials, which were repeated if necessary, and a 5-point calibration and validation procedure. Trials started with a fixation cross shown for 2,000 ms, followed by a prearrow display introducing the four pictures for 1,000 ms (see Figure 3). After that, an arrow indicated the target picture by pointing at one of the four pictures and disappeared after 500 ms. The four pictures remained on the screen for 2,000 ms until a gray, visual noise screen was presented as a cue for linguistic production (see Figure 3). Participants' task was to describe (i.e., speak/sign) the picture at which the arrow was

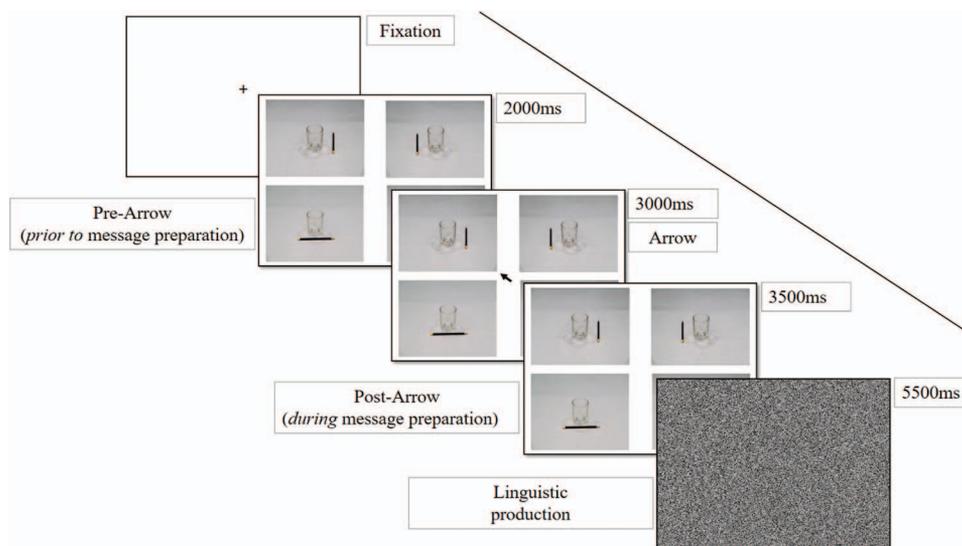


Figure 3. Timeline of trial structure. See the online article for the color version of this figure.

pointing (i.e., the target picture), to a confederate when the gray screen appeared. After each description, the confederate pretended to choose the described picture out of four pictures on a tablet. We instructed participants in such a general manner, because the confederates' cover task provided the participants with enough indications of what is expected from them. Furthermore, with this general instruction we aimed to allow unbiased linguistic descriptions as well as individual scanning of the display that is guided by the way participants will describe the pictures. The confederate was present to elicit a natural conversation and the confederates' hearing status was always matched to that of the participant (i.e., a deaf confederate for deaf participants, a hearing confederate for hearing participants). Importantly, participants were told that confederates were naive participants and the confederate was always another person than the experimenter to assure that the picture descriptions to the addressee are as informative as in natural conversations. Participants never questioned the status of the addressee during or after the experiment. All speakers were tested with the same hearing confederate and all signers were paired with the same deaf confederate. Participants initiated the next trial by pressing a button after they had described the target picture. Importantly, participants did not receive feedback regarding their spatial descriptions.

After this eye-tracking task, participants took part in a digital Corsi Block Tapping task to control individual differences in spatial memory. The eye-tracking task and the Corsi Block Tapping task were presented on an SMI RED-250 mobile laptop. The software package Presentation NBS 16.4 (Neurobehavioral Systems, Albany, CA) was used to control the eye tracker, present the stimuli, and register button presses during each trial. Triggers were sent from this software to the eye tracker. Eye-gaze was recorded binocular at a rate of 250 Hz (every 4 ms). Instructions were always given orally/visually in form of a video. A language background questionnaire was given at the end of the session to assess language use, language proficiency, deafness in family, and so forth. In total, the experimental session lasted approximately 45 min.

### Data Analysis

In this section, we will first describe how we analyzed signers' and speakers' spatial descriptions of the target pictures during the eye-tracking task. Next, we describe the analysis of the eye-gaze data first prior to message preparation followed by during message preparation. Finally, we will combine the two types of data (i.e., linguistic productions and eye gaze) by linking signer's linguistic productions to their eye-gaze competition during message preparation.

**Linguistic productions.** To assess the frequency of certain types of spatial encodings used by signers and speakers, we coded descriptions of the experimental items to describe left/right spatial configurations across more complex and less complex displays. In order to do this, we used ELAN, a free annotation tool (<http://tla.mpi.nl/tools/tla-tools/elan/>) for multimedia resources developed by the Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). Trained, native annotators (i.e., native Dutch annotator for Dutch data, deaf native NGT annotator for NGT data) performed annotations and codings of the data. All codings were checked by an additional coder to find consensus. If no consensus could be found, the trial was excluded from further analyses (3.61% of the data).

For speakers' picture descriptions in Dutch, we distinguished between two types of spatial encodings: categorical constructions (i.e., specific spatial relation noun left *links* or right *rechts*) and other alternative noncategorical forms (i.e., next to *naast*). In addition to speech, we also coded speakers' cospeech gestures, albeit very few, during their picture descriptions and coded per description whether any additional information was encoded iconically with the hands (i.e., about the object properties or spatial relation) or not.

For signers' spatial descriptions in NGT, we categorized linguistic strategies into two main groups depending on their type of iconicity. First, CLs (Figure 1A), which resemble visually space and object properties. Second, RLs (Figure 1B and Figure 1C),

which resemble space only. In addition, we coded for double strategies, thus descriptions that contained both RLs followed by CLs within one response, resembling space and additionally space and object properties (see Figure 4; CLs followed by RLs can also occur, although less frequently; see also Sümer, 2015 for TID).

This coded data was analyzed in R (Version 3.3.1; R Core Team, 2013). We used a logistic regression analysis with binomial link separately for speakers and signers. For speakers, we compared categorical (left/right) versus noncategorical (next to) strategies (Spatial Encoding Type) across more complex and less complex displays (Display Complexity) by using numeric contrasts (Helmert contrast). That is, for Spatial Encoding Type we coded categorical as  $+1/2$  and noncategorical as  $-1/2$ . For Display Complexity, we coded more complex displays as  $+1/2$  and less complex displays as  $-1/2$ . Similarly, for signers we compared linguistic strategies such as CLs and RLs (Spatial Encoding Type) across more complex and less complex displays (Display Complexity) by using numeric contrasts (Helmert contrast). That is, for Spatial Encoding Type we coded one the first level double strategies as  $-2/3$ , CLs as  $+1/3$ , and RLs as  $+1/3$ . On the second level we coded CLs as  $+1/2$  and RLs as  $-1/2$ . For Display Complexity, we coded more complex displays as  $+1/2$  and less complex displays as  $-1/2$ .

**Eye-gaze data.** For each trial, eye movements were recorded from prearrow onset (0 ms) until the four-picture display disappeared (3,500 ms). We analyzed fixation proportions (right eye only) across 50 ms continuous time bins for two specific time windows: the prearrow window and postarrow window. The prearrow window initiated immediately after presenting the fixation cross (0 ms) until arrow onset (1,000 ms; see Figure 3), thus participants did not know yet which spatial relation to describe. The prearrow window served as a baseline prior to message preparation (related to spatial relation encoding) to ensure that different aspects of our visual displays did not differ in their tendencies to attract gaze. The postarrow window initiated directly after arrow offset (1,500 ms) until production onset (3,500 ms; see Figure 3) capturing participants' linguistic planning phase linked to relational encoding. Thus, in the postarrow window we were able to assess speakers' and signers' viewing during message preparation (for relational encoding) over time.

We defined five different square-shaped Areas of Interest (AoI): one for each picture and one for the arrow region. We kept the size and positions of the five AoIs equal across all trials. Fixation data

were preprocessed and analyzed in R (Version 3.3.1; R Core Team, 2013). First, it was determined for each participant whether a fixation fell into a particular AoI in each of 70 time bins of 50 ms. Participants with more than 35% track loss across all trials were excluded from the analysis ( $N = 1$ ). Additionally, we excluded trials in which track loss was higher than 50% (4.55%).

To examine the eye-gaze patterns, we analyzed the fixation data by using linear mixed-effects regression models (Baayen, Davidson, & Bates, 2008) using the packages *lme4* (Version 1.1–19; Bates, Mächler, Bolker, & Walker, 2015) and *lmerTest* (Version 3.0–1; Kuznetsova, Brockhoff, & Christensen, 2017) to retrieve  $p$  values. We used the package *emmeans* (Length, 2019; Searle, Speed, & Milliken, 1980) to interpret significant interactions in these models.

Fixation proportions were corrected in both time windows for 200 ms to plan a first saccade (Matin, Shao, & Boff, 1993). We logarithmically transformed mean fixation proportions and time bins, which mathematically required shifting zeros and ones by  $+1.01$ .

**Eye-gaze competition prior to message preparation in speakers versus signers (prearrow window).** First, we wanted to ensure that signers and speakers did not differ from each other in their eye-gaze patterns before they knew which relation they would describe and that our displays did not have intrinsic features that guided attention. For more complex displays, Group (categorical predictor: speakers vs. signers), Bin (continuous predictor: for each 50 ms bin) and Spatial Relation Type (categorical predictor: lateral vs. sagittal vs. topological), were entered as fixed effects (predictors). Group was coded as a numeric contrast (Helmert contrast), that is, speakers as  $-1/2$  and signers as  $+1/2$ . Spatial Relation Type was also coded as a numeric contrast (Helmert contrast). Due to the presence of two lateral configurations (left and right), fixations across these two pictures were averaged. On the first level, we compared fixations to topological relations versus the averaged lateral and sagittal relations by coding topological as  $-2/3$ , lateral as  $+1/3$ , and sagittal as  $+1/3$ . On the second level we compared fixations to lateral versus sagittal relations by coding lateral as  $+1/2$  and sagittal as  $-1/2$ .

For less complex displays, the model was fitted with the same measurements and predictors as for the more complex displays and we applied the same Helmert contrast coding. Due to the absence of a second lateral relation but inclusion of a second topological

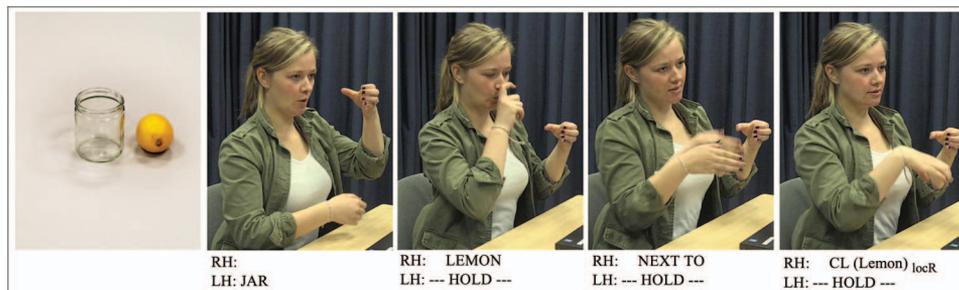


Figure 4. Example of a double strategy for “lemon is to the right of the jar” in NGT using both RL (next to) and CL within one description. These images are used with permission. See the online article for the color version of this figure.

relation the fixations for both topological relations (in and on) were averaged.

**Eye-gaze competition during message preparation in speakers versus (postarrow window).** To analyze visual attention to left/right spatial relations during message preparation (i.e., postarrow window) in more complex displays, we created three measures that assessed different types of eye-gaze competition: (a) mean of looks to target minus symmetrical competitor (*symmetrical competition*), (b) mean of looks to target minus sagittal competitor (*sagittal competition*), (c) mean of looks to target minus topological distractor (*topological competition*). In the model, Group (categorical predictor: speakers vs. signers), Bin (continuous predictor: for each 50 ms bin), and Competition Type (categorical predictor: symmetrical competition vs. sagittal competition vs. topological competition) were entered as fixed effects (predictors). As for the prearrow window, Group was coded as a numeric contrast, that is, speakers as  $-1/2$  and signers as  $+1/2$ . Competition Type was also coded as a numeric contrast. On the first level, we compared competition from the topological distractor versus competition from the other pictures by coding topological competition as  $-2/3$ , symmetrical competition as  $+1/3$ , and sagittal competition as  $+1/3$ . On the second level, we compared competition from the symmetrical competitor (coded as  $+1/2$ ) versus competition from the sagittal competitor (coded as  $-1/2$ ).

For less complex displays, the model was fitted with the same measurements and predictors as for the more complex displays, except for the symmetrical competition measurement, due to the absence of the symmetrical competitor. As in the above models, categorical predictors were coded as numeric contrast, however, for Competition Type, topological competition was coded as  $-1/2$  and sagittal competition as  $+1/2$ .

**The link between type of iconicity in linguistic production and eye-gaze competition during message preparation in signers only (postarrow window).** To further assess whether signers' eye-gaze competition can be predicted by the type of iconicity in their spatial encodings, we linked signers' linguistic productions to their eye gaze data during message preparation in more complex and less complex displays. For this, we excluded descriptions that could not be easily divided between types of iconicity (i.e., double strategies). Thus, we excluded double strategies (44.44% in complex displays and 41.97% in less complex displays) and consequently included only descriptions that contained only CLs or RLs. In the more complex display model, type of iconicity (categorical predictor: object + space resemblance vs. space resemblance), competition type (categorical predictor: symmetrical competition vs. sagittal competition vs. topological competition) and bin (continuous predictor: for each 50 ms bin) were entered as fixed effects (predictors). Type of iconicity was coded as numeric contrast, that is, visual resemblance as  $+1/2$  and semantic relevance as  $-1/2$ . Competition type was coded as numeric contrast (see above for eye-gaze competition model). For less complex displays, the model was fitted with the same measurements and predictors as for the more complex displays, except for the symmetrical competition measurement, due to the absence of the symmetrical competitor.

For all models, we conducted a backward selection procedure in which insignificant predictors were removed to obtain the most parsimonious model. The maximal random effects structure that converged in the model was implemented, which included random

intercepts for participants and items, as well as random slopes for group by items.

## Results

In this section we will first report the linguistic production data to assess the most frequently used linguistic forms by speakers and signers across more complex and less complex displays, including the type of iconicity in signers' encodings. After this, we will report the eye-gaze data beginning with the results prior to message preparation separately for more complex and less complex displays. Next, we will report results from eye-gaze competition during message preparation with the distinct measures separately for more complex and less complex displays. Finally, we will combine the two types of data (i.e., type of iconicity in linguistic productions and eye-gaze) and link signers' linguistic productions to their eye-gaze data across more complex and less complex displays.

### Linguistic Production

Figure 5 shows the most frequently used linguistic forms by speakers and signers during their picture descriptions (see Figure 5A for speaker data; see Figure 5B for signer data). For speakers, we investigated the frequency of categorical and noncategorical picture descriptions as a binary factor across the more complex and less complex displays. To investigate whether there are differences in speakers' descriptions, a logistic regression analysis yielded a main effect of Spatial Encoding Type ( $\beta = 0.094$ ,  $SE = -0.581$ ,  $z = 15.573$ ,  $p < .001$ ). This main effect suggests a preference of categorical spatial encodings (98.92% across trials) over noncategorical spatial encodings (1.08% across trials; see Figure 5A). There was no main effect of Display Complexity ( $\beta = 0.001$ ,  $SE = 0.581$ ,  $z = 0$ ,  $p > .1$ ) and no significant interaction between Display Complexity and Spatial Encoding Type ( $\beta = -0.022$ ,  $SE = 0.016$ ,  $z = -0.019$ ,  $p = .985$ ).

In addition to speech, we coded speakers' cospeech gestures during their picture descriptions to assess whether any additional information was encoded with the hands (i.e., about the object properties or spatial relation). Results indicate a small amount of cospeech gestures (13% in total). In all those instances, the cospeech gestures did not add any extra information about the location nor the properties of the figure and ground object. Rather, they indicated roughly the direction of the figure independent of the ground, which was already encoded in speech (e.g., small head or hand movement to the left/right of the speakers' body).

Furthermore, we assessed the frequency of linguistic strategies varying in their type of iconicity in signers. To investigate whether there are differences in signers' descriptions, a logistic regression analysis yielded a main effect of Spatial Encoding Type (Comparison 1<sub>[double vs. CL + RL]</sub>:  $\beta = -0.742$ ,  $SE = 0.054$ ,  $z = -13.668$ ,  $p < .001$ , Comparison 2<sub>[CL vs. RL]</sub>): ( $\beta = 0.816$ ,  $SE = 0.138$ ,  $z = 5.890$ ,  $p < .001$ ). The first comparison suggests that double strategies (43.22% across trials) are used more compared to using the average of RLs only and CLs only (see Figure 5B). In addition, the second comparison reveals that using CLs only (36.12% across trials) is the more preferred strategy than using RLs only (20.17% across trials). There was no main effect of Display Complexity

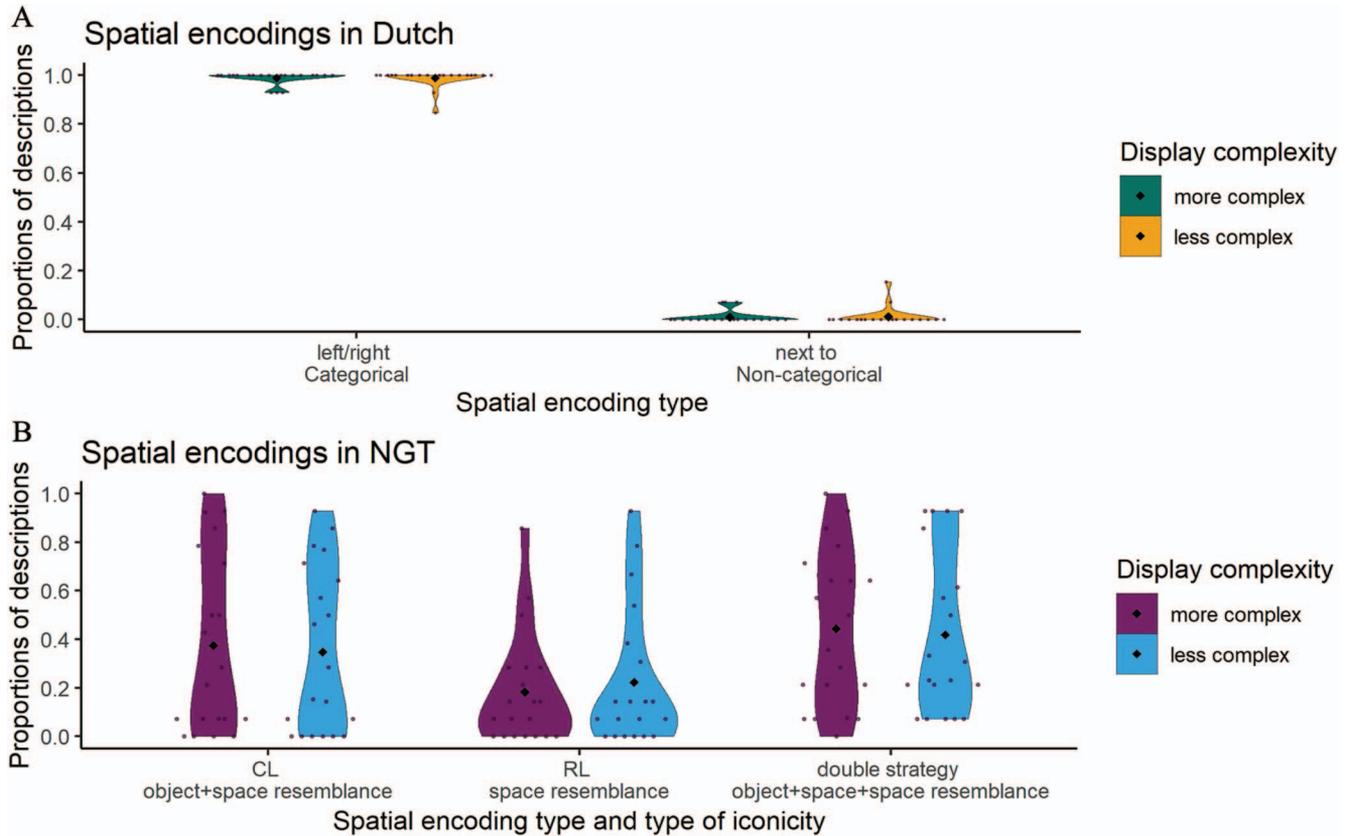


Figure 5. Types of spatial encodings used by Dutch speakers (A) and type of spatial encodings and iconicity in NGT signers' encodings (B). Rectangles represent the mean. Dots represent each data point (participant). The width of the violins represents the data distribution's density; the length of the violins depicts the range of data points. NGT = *Nederlandse Gebarentaal*. See the online article for the color version of this figure.

( $\beta = 0.009$ ,  $SE = 0.109$ ,  $z = -0.086$ ,  $p = .931$ ) and no interaction between Display Complexity and Spatial Encoding (Comparison 1<sub>[double vs. CL + RL]</sub>;  $\beta = -0.165$ ,  $SE = 0.221$ ,  $z = -0.749$ ,  $p = .454$ , Comparison 2<sub>[CL vs. RL]</sub>;  $\beta = 0.323$ ,  $SE = 0.277$ ,  $z = 1.166$ ,  $p = .244$ ).

### Eye-Gaze Patterns Prior to Message Preparation in Speakers Versus Signers (Prearrow Window)

**More complex displays.** First, we assessed whether signers and speakers did not differ from each other in their eye-gaze patterns before they knew which relation they would encode and that our displays did not have intrinsic features that guided attention in complex displays. The most parsimonious model incorporated two parameters for the random-effects structure of the data: random intercepts for participants and for the by-items random slope for Group. The model contained one main effect (Spatial Relation Type<sub>[lateral vs. sagittal]</sub>), two two-way interactions (Spatial Relation Type<sub>[lateral vs. sagittal]</sub>\*Group; Spatial Relation Type<sub>[lateral vs. sagittal]</sub>\*Bin) but no significant three-way interaction (Group\*Bin\*AoI; Figure 6). The first interaction with group suggests that while both speakers and signers look at lateral relations more the difference between the two spatial relation types is bigger for speakers than for signers indepen-

dent of the time course of display viewing. The second interaction with bin suggests increasing looks over time to lateral relations but not to sagittal relations independent of group. However the lack of an interaction between group and bin suggests that there is no difference between speakers and signers over time in the way they view the visual displays prior to message preparation (for relevant statistics and corresponding coefficients, see Table 2; for more information see Table S1 in the online supplementary materials, using the *emmeans* R package).

**Less complex displays.** For less complex displays, the most parsimonious model incorporated two parameters for the random-effects structure of the data: random intercepts for participants and for the by-items random slope for group. The model contained one main effect (Spatial Relation Type<sub>[lateral vs. sagittal]</sub>) and a two-way interaction (Spatial Relation Type<sub>[lateral vs. sagittal]</sub>\*Bin) but no significant three-way interaction (Group\*Bin\*AoI; Figure 7). The two-way interaction suggests, as in complex displays, increasing looks over time to lateral relations but not to sagittal relations, independent of group. The lack of an interaction with group and bin, as in complex displays, suggests that over time speakers and signers do not view the visual displays differently prior to message preparation (for relevant statistics and corre-

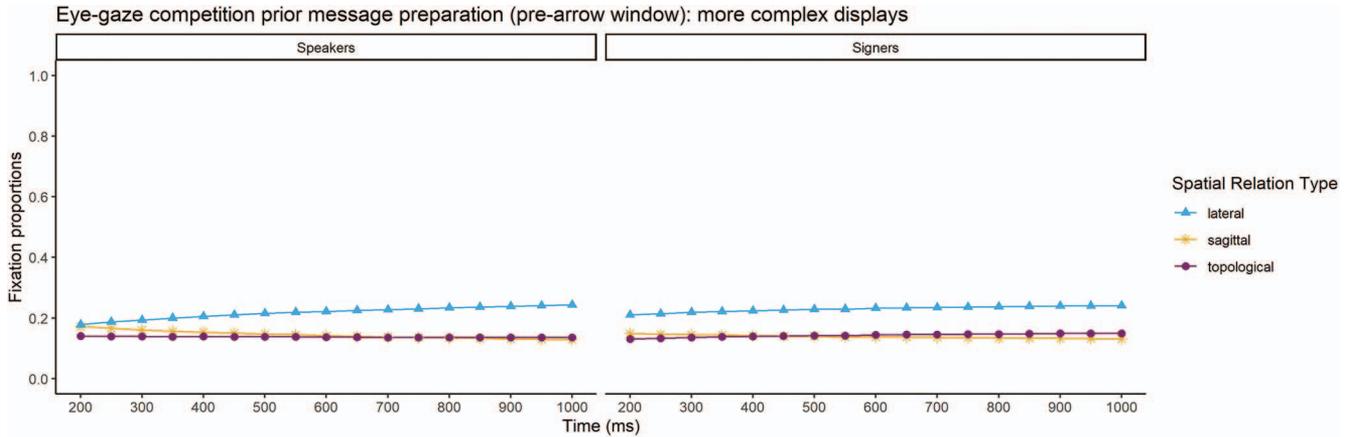


Figure 6. Eye-gaze competition across speakers (left panel) and signers (right panel) over time for more complex displays prior to message preparation (prearrow window). See the online article for the color version of this figure.

sponding coefficients, see Table 3; for more information see Table S2 in the online supplementary materials, using the *emmeans* R package).

### Eye-Gaze Competition During Message Preparation in Speakers Versus Signers (Postarrow Window)

**More complex displays.** We assessed visual attention to spatial relations during message preparation in complex displays. The most parsimonious model incorporated two parameters for the random-effects structure of the data: random intercepts for participants and for the by-items random slope for group. The model contained a main effect of group, bin and a two-way interaction (Group\*Bin). Thus, both groups experience competition equally from symmetrical competitors, sagittal competitors and topological distractors (see Figure 8; note that lower difference scores indicate more competition). The two-way interaction suggests, however, that signers but not speakers experience increased eye-gaze competition over time independent of the type of competition. The relevant statistics and corresponding coefficients are reported in Table 4 (for more

information see Table S3 in the online supplementary materials, using the *emmeans* R package).

**Less complex displays.** For less complex displays, the most parsimonious model incorporated two parameters for the random-effects structure of the data: random intercepts for participants and for the by-items random slope for group. The model contained three main effects (competition type, group; bin), three two-way interactions (Competition Type\*Group; Competition Type\*Bin, Group\*Bin) and a three-way interaction (Competition Type\*Group\*Bin). This three-way interaction suggests that both groups experience more competition from the sagittal competitor compared to the topological distractor (averaged over groups:  $\beta = -0.060$ ,  $SE = 0.003$ ,  $z = -17.520$ ,  $p < 0.001$ ; see Figure 9; note that lower difference scores indicate more competition). However, as in more complex displays, signers but not speakers experience increased eye-gaze competition over time. The relevant statistics and corresponding coefficients are reported in Table 5 (for more information see Table S4 in the online supplementary materials, using the *emmeans* R package).

Table 2  
Estimates, Standard Error, *t* Values, and *p* Values of the Main Effects of the Eye-Gaze Baseline Model for More Complex Displays for the Prearrow Window

Fixed effects	Estimate	Std. error	<i>t</i> value	<i>p</i> value
(Intercept)	0.061	0.009	6.430	.001
Relation type [other vs. topological]	0.015	0.019	0.757	.449
Relation type [lateral vs. sagittal]	-0.082	0.022	-3.678	.001
Group [speakers vs. signers]	-0.006	0.019	-0.295	.768
Bin	0.004	0.003	1.116	.264
Relation type [other vs. topological]:Group	0.040	0.039	1.044	.296
Relation type [lateral vs. sagittal]:Group	0.092	0.045	2.057	.040
Relation type [other vs. topological]:Bin	0.001	0.007	0.049	.961
Relation type [lateral vs. sagittal]:Bin	0.041	0.008	5.039	.001
Group [speakers vs. signers]:Bin	0.002	0.007	0.353	.724
Relation type [other vs. topological]:Group:Bin	-0.015	0.014	-1.078	.281
Relation type [lateral vs. sagittal]:Group:Bin	-0.031	0.016	-1.915	.055

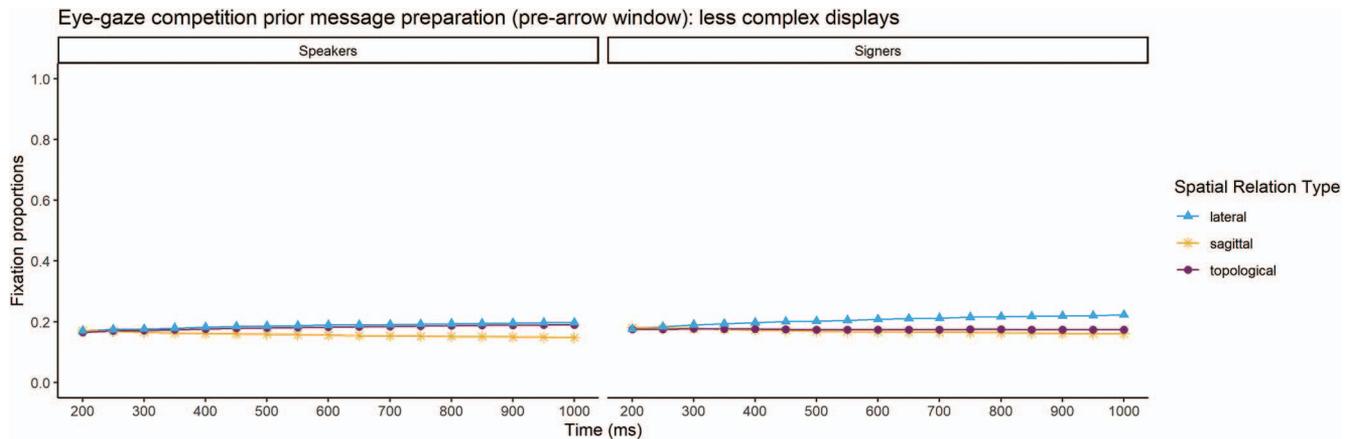


Figure 7. Eye-gaze competition across speakers (left panel) and signers (right panel) over time for less complex displays prior to message preparation (prearrow window). See the online article for the color version of this figure.

### The Link Between Linguistic Production and Eye-Gaze Competition During Message Preparation in Signers Only (Postarrow Window)

Finally, to assess whether signers' eye-gaze competition can be predicted by the type of iconicity in their spatial encodings, we linked the type of iconicity for each picture description (i.e., object + spatial resemblance vs. spatial resemblance) to signers' eye-gaze competition during the preparation of that description. For more complex displays, the most parsimonious model incorporated random intercepts for participants and for the by-items random slope for type of iconicity. The model contained three main effects (competition type<sub>[symmetrical vs. sagittal]</sub>\*type of iconicity, bin), four two-way interactions (Competition Type<sub>[other vs. topological]</sub>\*Type of Iconicity; Competition Type<sub>[symmetrical vs. sagittal]</sub>\*Type of Iconicity; Competition Type<sub>[symmetrical vs. sagittal]</sub>\*Bin; Type of Iconicity\*Bin), and two three-way interactions (Competition Type<sub>[other vs. topological]</sub>\*Type of Iconicity\*Bin; Competition Type<sub>[symmetrical vs. sagittal]</sub>\*Type of Iconicity\*Bin). The first comparison of this three-way interaction suggests that when planning CLs, signers experience more competition from sagittal and symmetrical competitors than from topological

distractors compared to planning RLs. In addition, the second comparison reveals that when signers plan RLs, they experience more competition over time from the symmetrical competitor than from the sagittal competitor compared to planning CLs (see Figure 10). The relevant statistics and corresponding coefficients are reported in Table 6 (for more information see Table S5 in the online supplementary materials, using the *emmeans* R package). Relating this to the results shown in Figure 8, signers in fact do divide their attention to the different competitors presented depending on the linguistic strategies they are preparing for. Thus, the current analysis on eye gaze linked to types of iconic forms reveals that differences in attention to the different competitors is attributed by iconicity in the relational encodings.

For less complex displays, the model did not contain a three-way interaction (Competition Type\*Type of Iconicity\*Bin;  $\beta = -0.244$ ,  $SE = 0.157$ ,  $t = -1.56$ ,  $p = .119$ ), suggesting that independent of the type of iconicity, signers experience more eye-gaze competition over time from the sagittal competitor compared to the topological distractor in less visually complex displays (for relevant statistics and corre-

Table 3  
Estimates, Standard Error, *t* Values, and *p* Values of the Main Effects of the Eye-Gaze Baseline Model for Less Complex Displays for the Prearrow Window

Fixed effects	Estimate	Std. error	<i>t</i> value	<i>p</i> value
(Intercept)	0.064	0.010	6.358	.001
Relation type [other vs. topological]	0.007	0.020	0.352	.725
Relation type [lateral vs. sagittal]	-0.071	0.024	-3.024	.002
Group	0.008	0.020	0.407	.684
Bin	0.004	0.004	1.164	.244
Relation type [other vs. topological]:Group	-0.044	0.041	-1.077	.282
Relation type [lateral vs. sagittal]:Group	-0.019	0.047	-0.403	.687
Relation type [other vs. topological]:Bin	-0.002	0.007	-0.324	.746
Relation type [lateral vs. sagittal]:Bin	0.031	0.009	3.547	.001
Group:Bin	-0.002	0.007	-0.275	.784
Relation type [other vs. topological]:Group:Bin	0.019	0.015	1.266	.206
Relation type [lateral vs. sagittal]:Group:Bin	0.008	0.017	0.454	.650

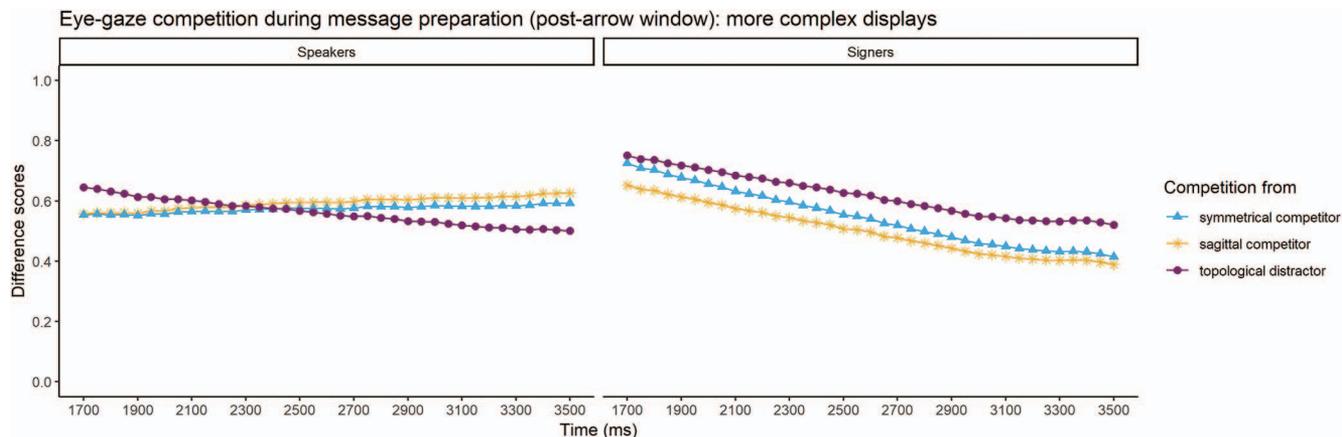


Figure 8. Eye-gaze competition across speakers (left panel) and signers (right panel) over time for more complex displays during message preparation (postarrow window). Smaller y-axis values indicate more eye-gaze competition. See the online article for the color version of this figure.

sponding coefficients, and graphs see Table S6 in the online supplementary materials).

### Discussion

To talk about space, spoken languages rely on arbitrary and categorical forms. In sign languages, however, the visual modality allows for visually motivated, that is iconic, form-to-meaning mappings of space onto the signing space. The aim of this study was to assess whether the iconic encodings about spatial relations in sign languages guide visual attention to spatial relations differently than the arbitrary and categorical spatial encodings in spoken languages during message preparation at the sentence level. We introduced an adapted version of the visual world eye-tracking experiment to investigate the link between message preparation (spoken vs. signed) and visual attention for spatial relations during message preparation. We expected differences in how visual attention is deployed during but not prior to message preparation, due to differences in the way speakers and signers linguistically encode spatial relations. Previous studies linking eye gaze to language production at the sentence level have focused only on spoken languages and have also found differences in eye gaze

in line with cross-linguistic patterning. This is the first study to investigate whether modality differences such as iconicity across sign and spoken languages also guide visual attention differently when planning messages at the sentence level.

Our study has several key findings. First, speakers and signers described spatial configurations in the most informative manner available (according to the expected modality preferences) regardless of the complexity of displays (i.e., presence vs. absence of a symmetrical competitor). That is, speakers used more categorical than noncategorical encodings while signers preferred iconic forms, CLs over RLs as well as double strategies (i.e., using both CLs and RLs within one description). Second, signers, but not speakers, experienced more eye-gaze competition over time from nontarget spatial configurations during message preparation (postarrow window). Importantly, prior to message preparation (prearrow window), eye-gaze patterns between speakers and signers were similar and did not indicate differences in eye gaze competition over time as we have observed during message preparation for relational encoding (i.e., postarrow window). Furthermore, signers' eye-gaze competition was predicted by the type of iconicity in signers' spatial encodings. These results together indicate

Table 4  
Estimates, Standard Error, *t* Values, and *p* Values of the Main and Interaction Effects of the Eye-Gaze Competition Model for More Complex Displays for the Postarrow Window

Fixed effects	Estimate	Std. error	<i>t</i> value	<i>p</i> value
(Intercept)	0.628	0.086	7.274	.001
Competition type [other vs. topological]	0.033	0.218	0.153	.878
Competition type [symmetrical vs. sagittal]	0.230	0.222	1.037	.300
Group	0.609	0.172	3.549	.001
Bin	-0.131	0.025	-5.302	.001
Competition type [other vs. topological]:Group	-0.632	0.436	-1.448	.148
Competition type [symmetrical vs. sagittal]:Group	-0.328	0.445	-0.737	.461
Competition type [other vs. topological]:Bin	-0.009	0.064	-0.137	.891
Competition type [symmetrical vs. sagittal]:Bin	-0.065	0.065	-0.993	.321
Group:Bin	-0.182	0.049	-3.697	.001
Competition type [other vs. topological]:Group:Bin	0.187	0.128	1.457	.145
Competition type [symmetrical vs. sagittal]:Group:Bin	0.108	0.131	0.825	.409

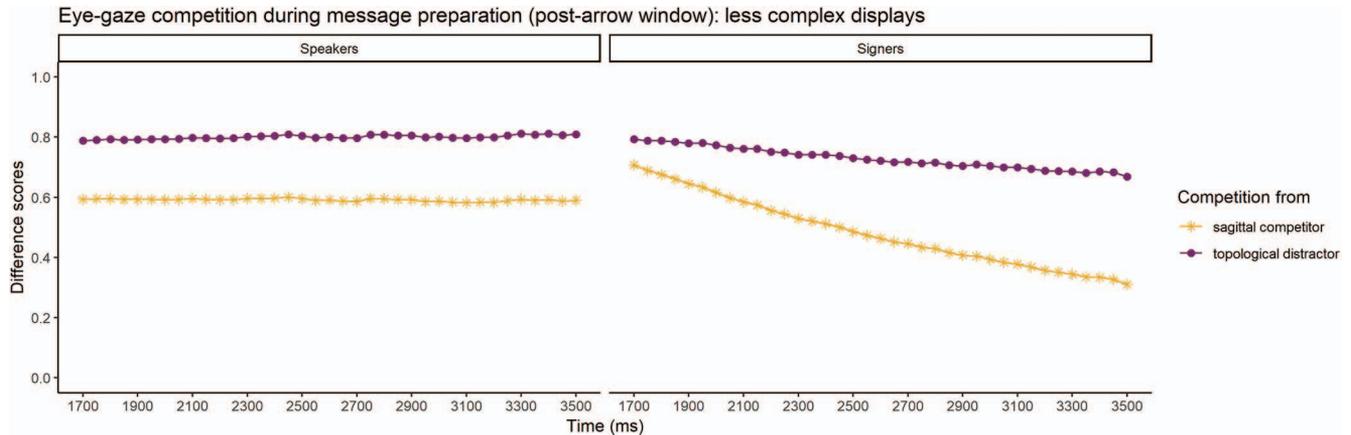


Figure 9. Eye-gaze competition across speakers (left panel) and signers (right panel) over time for less complex displays during message preparation (postarrow window). Smaller y-axis values indicate more eye-gaze competition. See the online article for the color version of this figure.

that the way sign languages are organized (i.e., iconic patterning) modulates signers' visual attention to spatial relations differently than that of speakers during message preparation. Below we will discuss the implications of our results in more detail.

### Speakers and Signers Prefer Modality-Specific Ways of Encoding Spatial Relations but Both Use the Most Informative Strategy to Describe Left/Right Spatial Relations

We determined to what extent Dutch speakers and NGT signers used different linguistic strategies to encode spatial relations across more or less complex displays. Speakers used more categorical than noncategorical encodings. Signers used iconic encodings and within those iconic preferences they favored CLs over using RLs, confirming previous findings on such a CL preference (Emmorey, 2002; Permiss, 2007; Sümer, 2015). Interestingly, NGT signers overall preferred to use double strategies (i.e., CLs and RLs within one description) compared to using one strategy alone. We are the first to demonstrate a preference in NGT for using both CLs and RLs within one description for the encoding of left/right relations.

Crucially, however, the complexity of displays influenced neither speakers' nor signers' linguistic strategies. That is, speakers did not produce more specific spatial relation nouns (e.g., left/right) to de-

scribe more complex displays and more general spatial relation nouns (e.g., *next to*) to describe less complex displays. Nor did signers use more double strategies or CLs in more complex displays compared to less complex displays. Instead, speakers and signers described spatial configurations in the most informative way available, regardless of the presence or absence of a symmetrical competitor. This is in line with previous research in which speakers were as informative about contrast objects (e.g., big vs. small sausage) during complex displays (eight picture grid) as well as simple displays (four picture grid; Davies & Kreysa, 2017). The current study demonstrated that these findings can be extended to even more complex displays (i.e.,  $2 \times 2$  grid with spatial relations) and to the visual signing modality.

### Signers, but not Speakers, Experienced More Eye-Gaze Competition From Spatial Configurations During Message Preparation (Postarrow Window) but not Prior to Message Preparation (Prearrow Window)

**Postarrow window (during message preparation).** For the postarrow window, which is where we expected message planning to occur specifically for relational encoding, speakers and signers showed competition equally from both the competitors and distractor. However, signers, but not speakers, experienced increasing eye-gaze competition from competitors and distractor over time. For speakers,

Table 5  
Estimates, Standard Error, *t* Values, and *p* Values of the Main and Interaction Effects of the Eye-Gaze Competition Model for Less Complex Displays for the Postarrow Window

Fixed effects	Estimate	Std. error	<i>t</i> value	<i>p</i> value
(Intercept)	0.611	0.066	9.300	.001
Competition type	0.423	0.127	3.328	.001
Group	0.727	0.131	5.536	.001
Bin	-0.118	0.019	-6.320	.001
Competition type:Group	0.823	0.254	3.235	.001
Competition type:Bin	-0.142	0.037	-3.804	.001
Group:Bin	-0.221	0.037	-5.914	.001
Competition type:Group:Bin	-0.246	0.075	-3.292	.001

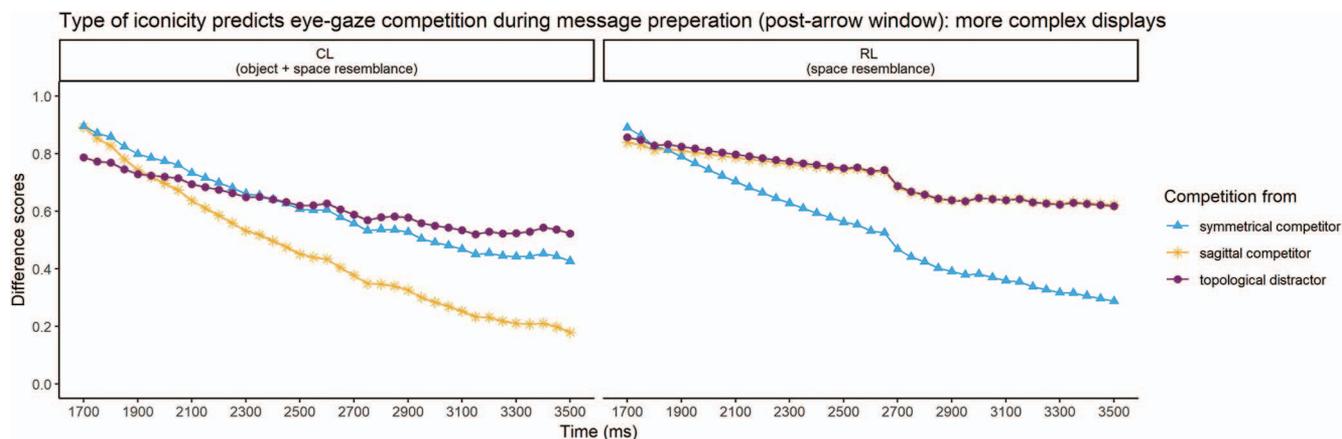


Figure 10. Type of iconicity predicts signers' eye-gaze competition in more complex displays during message preparation (left panel). Smaller y-axis values indicate more contrast competition. CL = classifier construction; RL = relation lexemes. See the online article for the color version of this figure.

planning categorical linguistic forms to describe spatial relations did not increase their eye-gaze competition. Signers, unlike speakers, have to map iconic information on to the signing space and thus need to identify the orientation, size, shape, and the relative locations of the objects in visually similar pictures. This seems to result in enhanced visual attention to spatial configurations, independent of whether it is visually or semantically related or not.

During viewing of less complex displays, similar differences arose in how speakers and signers allocate their visual attention. Thus, as for more complex displays, signers, but not speakers, experienced an increase in competition from spatial configurations during message preparation for less complex displays. Additionally, we found more competition from the sagittal competitor than from the topological distractor in less complex displays for both groups (which we did not observe in more complex-displays). This could be indeed due to fact that sagittal relations (require viewpoint) are perceived differently than topological relations (Levinson, 1996, 2003).

The reason that this difference between sagittal competitor and topological distractor only arises in less complex displays and not in more complex displays could be due to the visual complexity of displays (see Sorensen & Bailey, 2007, for evidence on display complexity overshadowing the eye movement–language processing link). Furthermore, it has been proposed that visual context such as shape similarity between presented stimuli within a visual world paradigm can constrain language-mediated anticipatory eye-movements (Hintz, Meyer, & Huettig, 2019). That is, the presence of too many competitors (and additionally the visual similarity between sagittal and lateral relations) and a distractor might have masked this difference in the complex displays. Furthermore, the use of different types of iconicity by signers might have also masked differences between types of spatial relations in complex displays in this analysis. As we will see below in the following section, eye-gaze patterns linked to type of iconicity in signers, in fact, reveal differences between sagittal and lateral compet-

Table 6  
Estimates, Standard Error, *t* Values, and *p* Values of the Main and Interaction Effects of the Model Assessing Whether the Type of Iconicity in NGT Spatial Encodings Predict Signers' Eye-Gaze Competition in More Complex Displays During Message Preparation (Postarrow Window)

Fixed effects	Estimate	Std. error	<i>t</i> value	<i>p</i> value
(Intercept)	1.412	0.143	9.855	.001
Competition type [other vs. topological]	-0.505	0.291	-1.734	.083
Competition type [symmetrical vs. sagittal]	-0.680	0.336	-2.024	.043
Type of iconicity	0.584	0.277	2.109	.035
Bin	-0.357	0.040	-8.822	.001
Competition type [other vs. topological]:Type of iconicity	1.272	0.582	2.184	.029
Competition type [symmetrical vs. sagittal]:Type of iconicity	-1.478	0.672	-2.199	.028
Competition type [other vs. topological]:Bin	0.153	0.086	1.785	.074
Competition type [symmetrical vs. sagittal]:Bin	0.207	0.099	2.101	.036
Type of iconicity:Bin	-0.190	0.081	-2.352	.019
Competition type [other vs. topological]:Type of iconicity:Bin	-0.395	0.171	-2.307	.021
Competition type [symmetrical vs. sagittal]:Type of iconicity:Bin	0.449	0.198	2.272	.023

Note. NGT = Nederlandse Gebarentaal.

itors versus topological distractors in the complex displays depending on which linguistic strategy they are preparing for (CLs vs. RLs).

**Prearrow window (prior to message preparation).** We had expected that during the prearrow window, signers' and speakers' viewing would not differ in ways linked to linguistic message preparation of spatial relations. Even though our results unexpectedly indicate a difference between signers and speakers, we argue that this difference is unrelated to the way signers and speakers prepare their linguistic message. In particular, speakers were more likely to prefer looking at lateral relations compared to signers, which was, however, independent of time (i.e., no interaction with bin, unlike what we see in the postarrow window). It appears that our visual displays create an overall preference to lateral relations for both groups, which seems to occur more strongly in more complex than in less complex displays. This preference might be rather due to the visual similarity between the two lateral relations and thus might be more driven by characteristics in the visual input (more strongly for speakers) but not related to message preparation (as there is no interaction with bin). Note that at this stage participants do not know which spatial relation they will encode and thus these looks are unlikely to be related to the relational encoding we are interested in. However, and most importantly, when signers and speakers concretely prepare their linguistic message (postarrow window), different patterns arise that progressed differently in time between signers and speakers while message preparation was unfolding (i.e., increasing competition for signers but not for speakers). Thus, once speakers and signers start to prepare the linguistic message after the appearance of the arrow (i.e., knowing which relation to describe), differences in visual attention arise that relate to the way they prepare their messages and these patterns appear to be different than what we can observe in the prearrow window (i.e., when participants do not know which spatial relation to describe). Nevertheless, we do not fully exclude the possibility that during the prearrow window participants engage in some sort of linguistic preparation for the lexical items; but importantly, we believe this does not occur for the relations between items. We argue that possible linguistic preparation ongoing in the prearrow window does not reflect the active preplanning stage of the concrete linguistic message regarding the spatial relations. Rather, the prearrow window might reflect attentional preferences based on visual similarity or dissimilarity of the pictures (as we indeed find a slight preference to look at lateral relations than at sagittal relations), but also by other factors, for instance covert object naming.

Note that such differences in the prearrow window might influence the interpretation of eye-gaze patterns in the postarrow window (see [Barr, 2008](#)). However, our results provide evidence against such a baseline influence. In particular, the differences between signers' and speakers' eye-gaze pattern in the prearrow window arise specifically for lateral relations over sagittal relations (i.e., more looks at lateral relations than sagittal relations for speakers compared to signers). However, in the postarrow window, the difference in visual attention between signers and speakers occurs across all competitor types and seems thus not to be a result of the group difference observed in the prearrow window.

## Signers' Visual Attention Was Predicted by the Type of Iconicity in Their Spatial Descriptions

Our data also revealed that in the postarrow window not only the modality differences of spatial language (arbitrary and categorical speech vs. iconic sign) influenced visual attention to spatial relations during message preparation, but also the type of iconicity within signers modulated their visual attention. That is, for more complex displays preparing to use CLs to describe left/right configurations led to eye-gaze competition from other spatial competitors compared to the planning of producing RLs. When planning to use CLs, signers experienced more eye-gaze competition from pictures that visually resembled the target picture, thus from the symmetrical and sagittal competitors. Planning CLs requires where to place the hands in space, thus, placing the figure handshape (e.g., long-elongated handshape for pen) to the sides of the ground handshape (e.g., left vs. right vs. front vs. behind the round handshape for cup). Consequently, this mapping of object shapes and spatial relation onto the signing space requires more visual attention to visually resembling spatial configurations. On the other hand, planning RLs resulted in enhanced competition only from pictures that were also semantically relevant, thus from the symmetrical competitor only. Planning RLs require the identification of the relative location in space, thus only semantically relevant pictures (i.e., symmetrical competitor) competed for visual attention. Competition from the topological distractor was not predicted by the type of iconicity as it is neither visually nor semantically relevant for describing left/right target configurations. This confirms our initial expectations about differences between viewpoint dependent spatial relations (sagittal and lateral) versus topological relations in complex displays when specific encodings are taken into account and justifies the choice of spatial relations used in our design.

For less complex displays, signer's visual attention was not predicted by the type of iconicity in their spatial descriptions. Due to the absence of a symmetrical competitor, only one visually similar picture was present in the display (sagittal competitor) and no picture was semantically relevant to the left/right target picture. Thus, signers experienced more competition from sagittal competitors than from topological distractors, independent of whether they planned CLs or RLs.

Taken together, the way signers planned to describe the target picture (i.e., CLs vs. RLs) shifted their attention toward different aspects of the visual display (visual resemblance vs. semantic relevance) that were relevant for message preparation. However, when a symmetrical competitor was absent (i.e., in less complex displays) the type of iconicity did not predict eye-gaze competition from the sagittal competitor or topological distractor. This finding again corroborates the above-mentioned finding that complexity of the visual display itself influences how visual attention is allocated for language planning.

## Conclusion

Overall, our results provide novel knowledge that message preparation for language production affects how speakers and signers allocate visual attention to spatial relations and thus suggests that "thinking for speaking" ([Slobin, 2003](#)) differs from "thinking for signing." We further show that sign production and

eye-gaze seem to be linked as speech production and eye gaze are linked at the sentence level (e.g., Griffin, 2004; Griffin & Bock, 2000; van de Velde et al., 2014), although further sign–eye-gaze links during actual signing need to be investigated.

Previous research has demonstrated that iconicity can facilitate (Grote & Linz, 2003; Thompson et al., 2009) or hinder (Thompson et al., 2010) sign language comprehension. Our findings reveal that iconicity can affect not only language comprehension but also visual attention during the planning of language production. Our findings show that iconicity effects can surpass the lexical level and can be extended to the sentence level. Furthermore, we extend previous evidence from neuroimaging studies showing that planning CLs not only activate different brain regions involved in encoding spatial language (e.g., parietal regions) than encodings with RLs (e.g., left hemisphere language regions; Emmorey et al., 2013), but can also affect visual attention during message preparation differently.

Overall, the current study provides evidence for the first time that the way languages encode information across modalities (arbitrary speech vs. iconic sign) influences visual attention to entities and relations between them during message preparation at the sentence level. It further suggests that the influence of language on visual attention goes beyond cross-linguistic differences found across spoken languages and that thinking for speaking might indeed differ from thinking for signing. This study opens up new avenues for looking at whether iconicity involved in different aspects of sign language (e.g., syntactic, discourse level) also influences visual attention. Finally, this study offers novel insights into the relationship between sign production and eye gaze by showing for the first time ever that the type of form–meaning mapping in a language can guide visual attention and can be taken further to the view that iconicity can bridge the gap between language and human experience.

## References

Arik, E. (2013). Space, time, and iconicity in Turkish sign language (TID). *Trames: Journal of the Humanities and Social Sciences*, 16, 345–358. <http://dx.doi.org/10.3176/tr.2012.4.03>

Assaneo, M. F., Nichols, J. I., & Trevisan, M. A. (2011). The anatomy of onomatopoeia. *PLoS ONE*, 6, e28317. <http://dx.doi.org/10.1371/journal.pone.0028317>

Bååth, R. (2014). *BayesianFirstAid: A package that implements Bayesian alternatives to the classical \*.test functions in R*. Retrieved from [http://www.sumsar.net/files/academia/baath\\_user14\\_abstract.pdf](http://www.sumsar.net/files/academia/baath_user14_abstract.pdf)

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412. <http://dx.doi.org/10.1016/j.jml.2007.12.005>

Barr, D. J. (2008). Analyzing ‘visual world’ eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59, 457–474. <http://dx.doi.org/10.1016/j.jml.2007.09.002>

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

Bellugi, U., & Klima, E. S. (1976). Two faces of sign: Iconic and abstract. *Annals of the New York Academy of Sciences*, 280, 514–538. <http://dx.doi.org/10.1111/j.1749-6632.1976.tb25514.x>

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, 1573–1581. <http://dx.doi.org/10.1037/a0020934>

Bunger, A., Skordos, D., Trueswell, J. C., & Papafragou, A. (2016). How children and adults encode causative events cross-linguistically: Implications for language production and attention. *Language, Cognition and Neuroscience*, 31, 1015–1037. <http://dx.doi.org/10.1080/23273798.2016.1175649>

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, 6, 84–107. [http://dx.doi.org/10.1016/0010-0285\(74\)90005-X](http://dx.doi.org/10.1016/0010-0285(74)90005-X)

Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain* (Unpublished doctoral dissertation). McGill University, Montreal, Canada.

Davies, C., & Kreysa, H. (2017). Looking at a contrast object before speaking boosts referential informativeness, but is not essential. *Acta Psychologica*, 178(Suppl C), 87–99. <http://dx.doi.org/10.1016/j.actpsy.2017.06.001>

Davies, C., & Kreysa, H. (2018). Look before you speak: Children’s integration of visual information into informative referring expressions. *Journal of Child Language*, 45, 1116–1143. <http://dx.doi.org/10.1017/S0305000918000120>

Dingemans, M. (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass*, 6, 654–672. <http://dx.doi.org/10.1002/lnc3.361>

Dingemans, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, 19, 603–615. <http://dx.doi.org/10.1016/j.tics.2015.07.013>

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* (pp. 171–209). Cambridge, MA: MIT Press.

Emmorey, K. (2002). The effects of modality on spatial language: How signers and speakers talk about space. In D. Quinto-Pozos, K. Cormier, & R. P. Meier (Eds.), *Modality and structure in signed and spoken languages* (pp. 405–421). New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511486777.019>

Emmorey, K., & Herzig, M. (2003). Categorical versus gradient properties of classifier constructions in ASL. In K. Emmorey (Ed.), *Perspectives on classifier constructions in signed languages* (pp. 222–246). Mahwah, NJ: Lawrence Erlbaum.

Emmorey, K., McCullough, S., Mehta, S., Ponto, L. L., & Grabowski, T. J. (2013). The biology of linguistic expression impacts neural correlates for spatial language. *Journal of Cognitive Neuroscience*, 25, 517–533. [http://dx.doi.org/10.1162/jocn\\_a\\_00339](http://dx.doi.org/10.1162/jocn_a_00339)

Emmorey, K. (2014). Iconicity as structure mapping. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 369, 20130301. <http://dx.doi.org/10.1098/rstb.2013.0301>

Flecken, M., Carroll, M., Weimar, K., & Von Stutterheim, C. (2015). Driving along the road or heading for the village? Conceptual differences underlying motion event encoding in French, German, and French–German L2 users. *Modern Language Journal*, 99(S1), 100–122. <http://dx.doi.org/10.1111/j.1540-4781.2015.12181.x>

Flecken, M., Von Stutterheim, C., & Carroll, M. (2014). Grammatical aspect influences motion event perception: Findings from a cross-linguistic non-verbal recognition task. *Language and Cognition*, 6, 45–78.

Goller, F., Lee, D., Ansorge, U., & Choi, S. (2017). Effects of language background on gaze behavior: A crosslinguistic comparison between Korean and German speakers. *Advances in Cognitive Psychology*, 13, 267–279. <http://dx.doi.org/10.5709/acp-0227-z>

Griffin, Z. M. (2004). Why look? Reasons for eye movements related to

- language production. In J. Henderson & F. Ferreira (Eds.), *The integration of language, vision, and action: Eye movements and the visual world* (pp. 213–248). New York, NY: Psychology Press.
- Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11*, 274–279. <http://dx.doi.org/10.1111/1467-9280.00255>
- Grote, K., & Linz, E. (2003). The influence of sign language iconicity on semantic conceptualization. In W. G. Müller & O. Fisher (Eds.), *From sign to signing: Iconicity in language and literature* (pp. 23–40). Amsterdam, the Netherlands: John Benjamins. <http://dx.doi.org/10.1075/ill.3.05gro>
- Hintz, F., Meyer, A. S., & Huettig, F. (2019). Visual context constrains language-mediated anticipatory eye movements. *Quarterly Journal of Experimental Psychology, 73*, 458–467. <http://dx.doi.org/10.1177/1747021819881615>
- Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology, 73*, 1–40. <http://dx.doi.org/10.1016/j.cogpsych.2014.04.001>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*, 1–26. <http://dx.doi.org/10.18637/jss.v082.i13>
- Length, R. (2019). emmeans: Estimated marginal means, aka least-squares means [Computer software]. Retrieved from <https://CRAN.R-project.org/package=emmeans>
- Levinson, S. C. (1996). Language and space. *Annual Review of Anthropology, 25*, 353–382. <http://dx.doi.org/10.1146/annurev.anthro.25.1.353>
- Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity*. New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511613609>
- Lieberman, A. M., Borovsky, A., Hatrak, M., & Mayberry, R. I. (2015). Real-time processing of ASL signs: Delayed first language acquisition affects organization of the mental lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 1130–1139. <http://dx.doi.org/10.1037/xlm0000088>
- Lieberman, A. M., Borovsky, A., & Mayberry, R. I. (2018). Prediction in a visual language: Real-time sentence processing in American Sign Language across development. *Language, Cognition and Neuroscience, 33*, 387–401. <http://dx.doi.org/10.1080/23273798.2017.1411961>
- Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception & Psychophysics, 53*, 372–380. <http://dx.doi.org/10.3758/BF03206780>
- Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66*, B25–B33. [http://dx.doi.org/10.1016/S0010-0277\(98\)00009-2](http://dx.doi.org/10.1016/S0010-0277(98)00009-2)
- Navarrete, E., Peressotti, F., Lerose, L., & Miozzo, M. (2017). Activation cascading in sign production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 302–318. <http://dx.doi.org/10.1037/xlm0000312>
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition, 108*, 155–184. <http://dx.doi.org/10.1016/j.cognition.2008.02.007>
- Perniss, P. (2007). *Space and iconicity in German sign language (DGS)* (Doctoral thesis, Radboud University, Nijmegen, the Netherlands). Retrieved from <http://dx.doi.org/10.17617/2.57482>
- Perniss, P., Özyürek, A., & Morgan, G. (2015). The influence of the visual modality on language structure and conventionalization: Insights from sign language and gesture. *Topics in Cognitive Science, 7*, 2–11. <http://dx.doi.org/10.1111/tops.12127>
- 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*, 227. <http://dx.doi.org/10.3389/fpsyg.2010.00227>
- Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: From a world of experience to the experience of language. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 369*, 20130300. <http://dx.doi.org/10.1098/rstb.2013.0300>
- Perniss, P., Zwitserlood, I., & Özyürek, A. (2015). Does space structure spatial language?: A comparison of spatial expression across sign languages. *Language, 91*, 611–641. <http://dx.doi.org/10.1353/lan.2015.0041>
- R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population marginal means in the linear model: An alternative to least squares means. *The American Statistician, 34*, 216–221.
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*, 109–147. [http://dx.doi.org/10.1016/S0010-0277\(99\)00025-6](http://dx.doi.org/10.1016/S0010-0277(99)00025-6)
- Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 157–191). Cambridge, MA: MIT Press.
- Sorensen, D. W., & Bailey, K. G. D. (2007). The world is too much: Effects of array size on the link between language comprehension and eye movements. *Visual Cognition, 15*, 112–115.
- Sümer, B. (2015). *Acquisition of spatial language by signing and speaking children: A comparison of Turkish Sign Language (TİD) and Turkish* (Unpublished doctoral dissertation). Radboud University Nijmegen, the Netherlands.
- Sümer, B., Perniss, P., Zwitserlood, I., & Özyürek, A. (2014). Learning to express left-right and front-behind in a sign versus spoken language. In P. Bello, M. Guarini, & M. McShane (Eds.), *Proceedings of the 36th annual meeting of the cognitive science society* (pp. 1550–1555). Austin, TX: Cognitive Science Society.
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description, Vol. 3: Grammatical categories and the lexicon* (pp. 225–282). Cambridge, UK: Cambridge University Press.
- Talmy, L. (2003). The representation of spatial structure in spoken and signed language. In K. Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 169–195). Mahwah, NJ: Erlbaum.
- Tanenhaus, M. K., & Trueswell, J. C. (2006). Eye movements and spoken language comprehension. In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 863–900). Oxford, UK: Elsevier. <http://dx.doi.org/10.1016/B978-012369374-7/50023-7>
- Thompson, R. L., Vinson, D. P., Fox, N., & Vigliocco, G. (2013). Is lexical access driven by temporal order or perceptual salience? Evidence from British Sign Language. *Proceedings of the Annual Meeting of the Cognitive Science Society, 35*, 1450–1455. Retrieved from <https://cogsci.mindmodeling.org/2013/papers/0273/paper0273.pdf>
- Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2009). The link between form and meaning in American Sign Language: Lexical processing effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 550–557. <http://dx.doi.org/10.1037/a0014547>
- Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2010). The link between form and meaning in British sign language: Effects of iconicity for phonological decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1017–1027. <http://dx.doi.org/10.1037/a0019339>
- 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*, 1443–1448. <http://dx.doi.org/10.1177/0956797612459763>
- Trueswell, J. C., & Papafragou, A. (2010). Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal*

- of Memory and Language*, 63, 64–82. <http://dx.doi.org/10.1016/j.jml.2010.02.006>
- van de Velde, M., Meyer, A. S., & Konopka, A. E. (2014). Message formulation and structural assembly: Describing “easy” and “hard” events with preferred and dispreferred syntactic structures. *Journal of Memory and Language*, 71, 124–144. <http://dx.doi.org/10.1016/j.jml.2013.11.001>
- 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 of London, Series B: Biological Sciences*, 369, 20130292. <http://dx.doi.org/10.1098/rstb.2013.0292>
- Vigliocco, G., Vinson, D. P., Woolfe, T., Dye, M. W., & Woll, B. (2005). Language and imagery: Effects of language modality. *Proceedings Biological Sciences*, 272, 1859–1863. <http://dx.doi.org/10.1098/rspb.2005.3169>
- Vinson, D., Thompson, R. L., Skinner, R., & Vigliocco, G. (2015). A faster path between meaning and form? Iconicity facilitates sign recognition and production in British Sign Language. *Journal of Memory and Language*, 82, 56–85. <http://dx.doi.org/10.1016/j.jml.2015.03.002>
- Wilson, K. M. (1999). Validity of global self-ratings of ESL speaking proficiency based on an FSI/ILR-referenced scale. *ETS Research Report Series*, 1999(2), i–64. <http://dx.doi.org/10.1002/j.2333-8504.1999.tb01811.x>
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In N. Calzolari, K. Choukri, A. Gangemi, B. Maegaard, J. Mariani, J. Odijk, & D. Tapias (Eds.), *Proceedings of the 5th international conference on language resources and evaluation* (pp. 1556–1559). Genoa, Italy: European Language Resources Association. Retrieved from [http://www.lrec-conf.org/proceedings/lrec2006/pdf/153\\_pdf.pdf](http://www.lrec-conf.org/proceedings/lrec2006/pdf/153_pdf.pdf)
- Zwitserslood, I. (2012). Classifiers. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language: An international handbook* (pp. 158–186). Berlin, Germany: De Gruyter Mouton. <http://dx.doi.org/10.1515/9783110261325.158>

Received September 18, 2018

Revision received March 4, 2020

Accepted March 4, 2020 ■

### E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <https://my.apa.org/portal/alerts/> and you will be notified by e-mail when issues of interest to you become available!