Does Space Structure Spatial Language?
Linguistic Encoding of Space in Sign Languages

Pamela Perniss (pamela.perniss@mpi.nl)
Inge Zwitserlood (inge.zwitserlood@mpi.nl)
Asli Özyürek (asli.ozyurek@mpi.nl)

Radboud University Nijmegen & Max Planck Institute for Psycholinguistics
PO Box 310, 6500 AH Nijmegen, Netherlands

Abstract
Spatial language in signed language is assumed to be shaped by affordances of the visual-spatial modality – where the use of the hands and space allow the mapping of spatial relationships in an iconic, analogue way – and thus to be similar across sign languages. In this study, we test assumptions regarding the modality-driven similarity of spatial language by comparing locative expressions (e.g., cup is on the table) in two unrelated sign languages, TID (Türk İşaret Dili, Turkish Sign Language) and DGS (Deutsche Gebärdensprache, German Sign Language) in a communicative, discourse context. Our results show that each sign language conventionalizes the structure of locative expressions in different ways, going beyond iconic and analogue representations, suggesting that the use of space to represent space does not uniformly and predictably drive spatial language in the visual-spatial modality. These results are important for our understanding of how language modality shapes the structure of language.

Keywords: iconicity; language modality; spatial language; locative expression; sign language

Introduction
Despite the difference in modality of expression, signed (visual-spatial) and spoken (vocal-aural) languages similarly conform to principles of grammatical structure and linguistic form (Klima & Bellugi, 1979; Liddell, 1980; Padden, 1983; Stokoe, 1960; Supalla, 1986). However, in signed language, the use of the hands as primary articulators within a visible spatial medium for expression (i.e. the space around the body) has special consequences for the expression of visual-spatial information (e.g. of referent size/shape, location, or motion).

Spatial language, such as locative expressions, is a primary domain in which modality affects the structure of representation. Locative expressions in both signed and spoken language are characterized by linguistic encoding of entities and the spatial relationship between them (cf. Talmy, 1985). However, sign language locative expressions differ radically from those in spoken language in affording a visual similarity (or iconicity) with the real-world scenes being represented. For example, a signed expression of the spatial relationship between a house and a bicycle is clearly iconic of the scene itself. In the example from American Sign Language (ASL) in Figure 1, the signer depicts a bicycle as being located beside a house by placing her hands (her left hand representing the house in still 2; her right hand representing the bicycle in still 4) next to each other in sign space. The spatial relationship between the signer’s hands represents the spatial relationship between the referents, whereby the handshapes are iconic with certain features of the referents (e.g. the inverted cupped hand to represent the bulk of a house). In contrast, there is no resemblance, or iconicity, between the actual scene and the linguistic form of a spoken language locative expression, as e.g. the English expression There is a bicycle next to the house.

In general, spoken languages exhibit a wide range of cross-linguistic variation in the encoding of spatial relationships in locative expressions, both in the devices used and in their morphosyntactic arrangement (Grinevald, 2006; Levinson & Wilkins, 2006). For example, spoken language locative expressions exhibit the use of adpositions, like the spatial prepositions used in English or the case-marking postpositions used in Turkish, or different types of locative or postural verbs (as in Ewe (Ghana) or Tzeltal (Mexico)).

Such variation is not expected in signed languages, however. Instead, signed languages are assumed to be structurally homogeneous in the expression of spatial relationships. The affordances of the visual-spatial modality for iconic, analogue spatial representation are assumed to be the primary force in shaping spatial expression, thus creating fundamental similarities in spatial language across different sign languages (e.g. Aronoff, Meir, Padden & Sandler, 2003; Emmorey, 2002). A consequence of this assumption of similarity, rooted in the notion that signers will exploit the iconic affordances of the modality where possible, has been a dearth of empirical investigation in this domain. Where the encoding of spatial relationships is mentioned in the literature, its iconic character is stated as fact, and conforms to the underlying assumption that spatial relationships will be represented in an iconic, analogue way,
as afforded by the modality (Morgan, Herman, Barriere & Woll, 2008; Pyers et al., 2010; Talmy, 2003).

In particular, the use of space allows the iconic, analogue representation (i.e. one-to-one mapping) of a real spatial configuration onto sign space, while the use of the hands allows the iconic representation of referents by reflecting salient object features in the handshape (as with the cupped hand for the house in Figure 1).

These iconic affordances of the modality for spatial expression are primarily exemplified in the system of so-called classifier predicates, common to essentially all sign languages studied to date (see Schembri, 2003 for an overview). In these morphologically complex predicates, the handshape classifies the referent either by depicting certain of its featural properties, in entity classifiers (e.g. a grasping hand to represent a cup) or by reflecting its manipulative functionality, in handling classifiers (e.g. a grasping hand to represent a cup). The placement of the hand in sign space encodes the location of the referent in real space. Furthermore, the use of both hands simultaneously visually encodes the spatial relationship between two referents. That is, through the relative simultaneous positioning of classifier predicates in sign space, a signer can give an iconic, analogue representation of relationships like The cup is on the table or The bicycle is next to the house, including metric information to encode the relative distance between the bicycle and the house. The system of classifier predicates in sign languages has been described as reflecting characteristics of visual scene parsing, making use of elements that are less categorical than the spatial devices of spoken languages (e.g., spatial prepositions such as ON, IN etc.), and which derive instead from the analogue conceptual representation of spatial scenes (Liddell, 2003; Talmy, 2003). As such, the structure of locative expressions in sign languages is taken to be straightforward, derivable from the iconic properties of the modality, and without the use of abstract spatial categories like in spoken languages.

However, extensive empirical investigation of spatial expressions within individual sign languages, and especially between different sign languages, has been sparse. Previous claims have been based on few examples from few signers, and have not generally been elicited within a communicative, discourse context. Moreover, there have been few direct quantitative and qualitative comparisons between sign languages using the same task. Thus, to date, there is not, in fact, enough evidence to know whether and to what extent modality-driven aspects of spatial description in signed language are really generalizable across sign languages.

In the present study, we test assumptions regarding the modality-driven, iconically-motivated form of spatial representation in signed language by comparing locative expressions between TID (Türk İşaret Dili, Turkish Sign Language) and DGS (Deutsche Gebärdensprache, German Sign Language), two historically unrelated sign languages, yet with historically and socio-linguistically comparable backgrounds (Deringil, 2002 for TID; Vogel, 1999 for DGS). We look at similarities and differences in the expression of spatial relationships between these two languages to investigate the extent to which the iconic and analogue affordances of representation in the visual-spatial modality are actually exploited in discourse-based spatial description.

If the affordances of the modality are the fundamental force in shaping locative expression, then we should expect to find overall similarity between the two sign languages. Specifically, we should expect to find the use of classifier predicates to iconically represent referents and their locations. We should also find the use of simultaneity – in particular, simultaneous classifier constructions – to directly encode the spatial relationship between referents in an analogue fashion. In addition, as a further effect of the visual modality, and as predicted by non-linguistic Gestalt principles, we should find that the (larger) Ground object is represented before the (smaller) Figure object, e.g. as is also common in drawing (Emmorey, 1996; Emmorey et al., 2002).

If, on the other hand, individual sign languages conventionalize less iconic/analogue and more abstract representations of spatial semantic notions, then we would expect to see language-specific differences between the sign languages (i.e. as we see in spoken languages).

Method

Participants

Twelve adult signers were recruited for each sign language (TID and DGS) from Deaf communities in Izmir (Turkey) and Aachen and Essen (Germany). All signers were deaf, and acquired sign language natively from birth from their deaf parents (all 12 TID signers; 9 DGS signers) or from other deaf adults and children through early exposure (3 DGS signers). Twelve additional deaf, adult signers were recruited for each sign language from the same communities as addressees.

Materials

Stimulus materials were photographs of objects in spatial relationships. Seven object types (cups, boats, cows, birds, plates, pens, paintings) served as Figures and were placed in ON relationships with a Ground (table, shelf, wall, or water) or NEXT-TO relationships with each other. Each Figure object type occurred in four different pictures, varying in number of tokens (1, 2, 3 or 4, many), as shown for cups in Figure 2 below. This yielded a total of 28 stimulus pictures.

Figure 2. Stimulus materials depicting ON (cup on table) and NEXT-TO (cup next to cup) spatial relationships between objects.
Procedure
Stimulus pictures were presented to signers one at a time on a laptop screen placed to their side. When signers had looked at a picture sufficiently, they clicked on an arrow at the bottom of the screen to make the screen white, turned to the addressee seated opposite, and gave a description of the picture. From sheets containing 24 thumbnail pictures at a time, including pictures in the stimulus set, addressees were asked to identify the picture described by the signer. Three different video recordings of all descriptions were made: (1) front view of signer; (2) front view of addressee; (3) top view of signer and addressee.

Coding
All descriptions were transcribed and coded using the multimedia linguistic annotation tool ELAN. Data was coded for the expression of ON and NEXT-TO spatial relationships by looking at: (1) order of mention of Figure(s) and Ground (for ON relationships); (2) type of localization device for Figure objects (e.g. classifier predicates); and (3) the use of simultaneous constructions to encode the Figure-Ground (for ON) or Figure-Figure (for NEXT-TO) spatial relationships.

Results
We analyzed descriptions that contained expression of the spatial relationships depicted in the stimulus pictures through the explicit mention of the Ground object (e.g. table) and localization of the Figure object(s) in relation to each other and to the Ground. The total number of locative expressions analyzed was 271 for TID and 287 for DGS.

Order of mention and type of localization device
We first investigated the order of mention of Figure and Ground objects between the two sign languages. We found that both TID and DGS signers are significantly more likely to express the Ground before the Figure in locative expressions (paired samples t-test DGS: \( t(11) = 23.35, p < .001 \), Ground before Figure; \( M = .96, SD = .07 \), Figure before Ground : \( M = .05, SD = .07 \); paired samples t-test TID: \( t(11) = 24.67, p < .001 \), Ground before Figure; \( M = .95, SD = .06 \), Figure before Ground : \( M = .05, SD = .06 \)).

We next looked at the use of localization devices in TID and DGS. As predicted by the assumption of maximum iconic similarity between form and meaning in signed locative expressions, we found a predominance of use of classifier predicates to represent and localize referents in both sign languages. However, there was a slight higher overall preference for the use of classifier predicates in DGS (72% of all locative expressions) compared to TID (63% of all locative expressions) that approached significance \( (F(1,22) = 2.78, p = .110) \). In addition to classifier predicates, we also found a range of other devices used in both sign languages, some of which occurred in both sign languages, but differed in overall preference of use, and some of which were specific to either DGS or TID, and have not been previously described in the literature.

For example, signers of both sign languages used signs that traced the outline of the Figure objects at a location in sign space. These signs, described as Size and Shape Specifiers (SASSes) in the literature (Supalla, 1986) were used in 7% of TID descriptions and in 19% of DGS descriptions, exhibiting a significant difference in overall preference of use \( (F(1,22) = 9.45, p < .05) \). The direct placement of lexical signs (used to name objects, as e.g. HOUSE in Figure 1) to localize referents in space (Emmorey, 1996) also occurred in both sign languages, but again differed significantly in overall preference, being used in 30% of TID descriptions and only 3% of DGS descriptions \( (F(1,22) = 57.65, p < .001) \).

We also found the use of language-specific devices, the use and function of which are notable with respect to the encoding of NEXT-TO spatial relationships and different means of simultaneous referent representation. Before we describe these forms in detail, we first present general results regarding the use of simultaneity for the representation of ON and NEXT-TO relationships.

Use of simultaneity
ON relationships In contrast to what is predicted by the iconic/analogue use of space to represent space, we found little use of simultaneous constructions to encode ON relationships (e.g. holding a classifier predicate representing a table, with one hand, while placing the other hand, as a classifier predicate representing a cup, on top of it to express the relationship of a cup on a table). Focusing first on descriptions of stimulus pictures that contained only one Figure object, we found that this strong iconic affordance of the modality was exploited in only 18% of one-object descriptions in TID and in 7% of one-object descriptions in DGS. The difference between the two sign languages approached significance \( (F(1,22) = 3.54, p = .073) \). The simultaneous representation of ON relationships in the remaining stimulus pictures, i.e. containing two or more Figure objects, was likewise very rare in both sign languages (5% of TID and 1% of DGS descriptions). Again, TID signers favored the use of simultaneity to represent spatial ON relationships compared to DGS signers, a difference which is significant in this comparison \( (F(1,22) = 4.60, p < .05) \). We are careful, however, to point out that the differences across sign languages in the use of simultaneous constructions to encode ON relationships should be interpreted with some caution, given low frequency.

NEXT-TO relationships Compared to the ON relationships, the description of stimulus pictures containing NEXT-TO relationships (e.g. the cups are next to each other) exhibited an overall greater use of simultaneity in both TID and DGS. Though the sign languages did not differ in the total amount of use of simultaneity (41% for TID, 47% for DGS; \( F(1,22) = .65, p = .429 \)), we found a very striking difference both in...
the type of simultaneous representation and in the iconic nature of the localizing devices between the sign languages.

Specifically, we found a striking difference between TID and DGS in the use of a unimanual vs. a bimanual device for the simultaneous representation of referent locations for NEXT-TO relationships. As shown in Figure 3, TID signers were significantly more likely than DGS signers to encode referents in a NEXT-TO arrangement ($F(1,22) = 27.58, p < .001$), using a unimanual form, in which each finger represents an entity, thereby simultaneously representing multiple referents (an example of which is shown in Figure 4a). DGS signers, on the other hand were significantly more likely than TID signers to use a bimanual form ($F(1,22) = 32.783, p < .001$), in which each hand represents a separate referent (an example of which is shown in Figure 4b).

![Figure 3. Use of unimanual and bimanual simultaneity in TID and DGS in encoding of NEXT-TO relationships.](Image)

**Figure 3.** Use of unimanual and bimanual simultaneity in TID and DGS in encoding of NEXT-TO relationships.

The TID unimanual form represents the semantic notion of "NEXT-TO-ness" (of multiple referents), but does so by simultaneously representing referents as being in a particular spatial arrangement – namely next to each other. This argument is motivated by the following factors. First of all, the unimanual form often appeared in conjunction with numerals as well as with other, more iconic referent localization devices (as in the example shown in Figure 5). We argue that the information encoded in these various forms used within a single description (i.e. the simultaneous unimanual form, numerals, and individual, sequential localizations of Figure objects) overlaps partially, but not fully, and that they exhibit important semantic differences.

Most simply, numerals obviously encode the specific number of referents, but do not localize referents or provide any featural information about them. The use of individual localization predicates (as in the direct localization of noun signs in still 1 of Figure 5, or as in classifier predicates) also encodes the specific number of referents, but does so by marking distinct locations in space with the hand, providing iconic information about both referent shape and location. Moreover, as was typical across TID signers, the signer in Figure 5 does not (in still 1) use simultaneity to localize the boats (in contrast to the DGS form shown in Figure 4b), such that the relative locations of the boats with respect to each other is encoded only through the sequential marking of distinct locations. Finally, the TID unimanual form that is in particular focus here also encodes the specific number of referents, but does so by simultaneously representing referents as being in a particular spatial arrangement – namely next to each other. We would not expect the use of this form for multiple objects that are not in a neat, side-by-side arrangement, arguing for a distinct semantics beyond encoding only number information.

FIGURE 5. Example of TID description containing localization of iconic sign for boat at four different locations, followed by numeral four, and then the unimanual NEXT-TO locative predicate to represent the four boats next to each other.

In sum, these characteristics of form and usage suggest that the most salient semantic function of this unimanual TID form is to encode the semantic notion of "NEXT-TO-ness" (of a specific number of entities). Other iconic mappings, i.e. representing featural information about the referents and more specific, metric configurational information (i.e.,
analogue mapping of spatial relations among entities), are lost in this locative device.

The language-specific bimanual device we found used by the DGS signers can be similarly characterized as a specific marker for the spatial semantic notion of NEXT-TO-ness. In this form, one hand holds at the initial referent location, while the other hand sequentially marks the locations of the other referents. The stationary hand functions as a visual anchor, thereby highlighting the NEXT-TO configuration of all the represented referents. Crucially, this type of anchored bimanual form often occurred with a generic, vertically-held flat handshape (as shown in Figure 4(b)), which could – similarly to the unimanual TİD form – be used for all object types. In this usage, with a generic handshape, the form exhibits an abstraction from the (afforded) iconic representation of the referents themselves. However, the bimanual anchor form, explicitly marking the semantic aspect of NEXT-TO-ness, was also frequently used with more iconic referent representations, in particular, with classifier predicates (e.g. a curved hand to represent a cup).

Finally, it is worth noting that the proportion of TİD bimanual encodings shown in Figure 3 result from a very different type of simultaneity than is present in the DGS bimanual construction shown in Figure 4(b). The TİD bimanual encodings were essentially all (99%) used to encode the spatial configuration between two Figure objects (e.g. two cups on the table), and resulted from the placement of both hands into sign space at the same time. DGS signers also used simultaneous placement of both hands in sign space, making up half of the DGS bimanual NEXT-TO encodings (56%). However, the remaining 44% of DGS bimanual NEXT-TO encodings used the anchored form, in which one hand keeps the encoded relation visually represented, while the other hand moves. In contrast, the anchor-hand strategy was almost entirely absent in TİD descriptions (1%).

**Discussion**

Overall, our results do not lend overwhelming support to the assumption of similarity of spatial language across signed languages as a result of the iconic affordances of the visual-spatial modality driving the shape of locative expression.

We found similarities between TİD and DGS in the preference for Ground before Figure encoding, as well as in the predominance of use of classifier predicates, which allow the iconic representation of features of the referents, as well as of the relative locations of referents in space. However, we also found crucial, and unpredicted, differences between TİD and DGS in the use of space to encode spatial relationships as well as less iconic and analogue ways of expressing spatial relations in each language.

Firstly, signers of both sign languages used localization devices other than classifier predicates in language-specific ways. Furthermore, contrary to what is predicted by the iconic, analogue affordances of the modality, we found an overall low occurrence of simultaneity to represent the spatial relationship between two (or more) referents – with a lower occurrence for ON relations than for NEXT-TO relations. Finally, and of particular importance, we found a striking difference in the way in which TİD and DGS signers used simultaneity to encode NEXT-TO relationships, going beyond iconicity and one-to-one analogue mapping of spatial relationships between entities, and favoring more abstract semantic relations (i.e. NEXT-TO-ness).

Our results make clear that the iconic affordances of visual-spatial modality do not straightforwardly lead to a certain way of encoding spatial relationships. Instead, different sign languages exhibit different preferences for encoding spatial relationships and devise language-specific forms for encoding certain spatial semantic aspects, which may be less iconic/analogue and more abstract. Thus, in spite of the iconic affordances of the visual-spatial modality in the spatial domain, a descriptive, usage-based analysis – as we have presented here – shows that individual sign languages do not exploit the possibilities of iconic/analogue representation in the same way nor to the same degree. Our study has shown that the visual-spatial modality may provide a stock of iconic affordances leading to more similarities between sign languages than between spoken languages, but that the exploitation and integration of these affordances into linguistic structure cannot be broadly generalized.

In addition, the difference in frequency of use of simultaneity in the encoding of ON and NEXT-TO relationships in both TİD and DGS points to another interesting similarity with patterns of locative encoding found in some spoken languages. The structure of locative expressions may exhibit sensitivity to semantic and pragmatic constraints, e.g. motivated by the prototypicality of spatial relationships. For example, in spoken Turkish, typical and expected spatial (as well as temporal) relationships are expressed with a very general locative case marker da/de (at) (as in example 1). The specific nature of an ON relationship (i.e. with a spatial noun specifically encoding ON-ness) need only be encoded for atypical, non-canonical, or unexpected spatial configurations (as in example 2).

(1) fíncan masa-da
cup    table-LOC
'The cup is on the table'

(2) ayakkabi masa-nun üst-ün -de
shoe   table-GEN top-POSS-LOC
'The shoe is on the table'

However, to express NEXT-TO relationships in spoken Turkish, the general locative da/de must always be further specified with spatial noun (as in example 3).

(3) kalem kağdın yan -m -da
pen    paper-GEN side-POSS-LOC
'The pen is next to the paper'

This suggests that NEXT-TO relationships are inherently considered to be less canonical and less predictable than ON relationships. In a similar way, ON relationships in both TİD
and DGS were expressed less explicitly, i.e. relying on less explicit analogue spatial representation through the use of simultaneity, than were NEXT-TO relationships.

In conclusion, our study demonstrates the importance of going beyond the idea of the iconic affordances provided by the visual-spatial modality (specifically, the use of space to represent space, and the use of the hands to represent referents) to fully understand spatial expression in sign languages. Our study contributes to a more comprehensive understanding of how modality contributes to language structure, opening up further avenues of research into the comparative study of sign languages in the domain of spatial language. Finally, our study highlights the importance of taking discourse and pragmatic motivations into account in explaining linguistic structure, in any modality and of being open to the notion of linguistic diversity, in both signed and spoken language modalities.

Acknowledgments
This work was supported by a NWO VIDI grant (#27670009) awarded to the last author. The first and second authors were supported as postdocs from these grants. We would like to thank Uwe Zelle and Hasan Huseyn Yılmaz for assistance with data collection and transcription.

References


