Spatial Language Use Predicts Spatial Memory of Children: Evidence from Sign, Speech, and Speech-plus-gesture

Dilay Z. Karadöller (dilay.karadoller@mpi.nl)
Centre for Language Studies, Radboud University, Nijmegen, The Netherlands
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
Wundtlaan 1, 6525XD, Nijmegen, The Netherlands

Beyza Sümer (b.sumer@uva.nl)
University of Amsterdam, Amsterdam, The Netherlands
P.C. Hoofthuis, Office: 6.40, Spuistraat 134, 1012 VB, Amsterdam, The Netherlands

Ercenur Ünal (ercenur.unal@ozyegin.edu.tr)
Ozyegin University, İstanbul, Turkey
Nişantepe Mahallesi, Orman Sokak, 34674, Çekmeköy, İstanbul, Turkey

Aslı Özyürek (asli.ozyurek@mpi.nl)
Centre for Language Studies & Donders Center for Cognition, Radboud University, Nijmegen, The Netherlands
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
Wundtlaan 1, 6525XD, Nijmegen, The Netherlands

Abstract

There is a strong relation between children's exposure to spatial terms and their later memory accuracy. In the current study, we tested whether the production of spatial terms by children themselves predicts memory accuracy and whether and how language modality of these encodings modulates memory accuracy differently. Hearing child speakers of Turkish and deaf child signers of Turkish Sign Language described pictures of objects in various spatial relations to each other and later tested for their memory accuracy of these pictures in a surprise memory task. We found that having described the spatial relation between the objects predicted better memory accuracy. However, the modality of these descriptions in sign, speech, or speech-plus-gesture did not reveal differences in memory accuracy. We discuss the implications of these findings for the relation between spatial language, memory, and the modality of encoding.

Keywords: spatial language; spatial memory; sign language; co-speech gesture

Introduction

Previous research shows strong relation between training children in spatial language and their later spatial memory accuracy (e.g., Gentner, 2016; Lowenstein & Gentner, 2005). However, this link has been established through studies that focus on speech alone and neglect sign languages and co-speech gestures (except for Gentner, Özyürek, Gürcanlı & Goldin-Meadow, 2013 that focused on deaf children prior to being exposed to a sign language). Visual-spatial forms of communication through sign or co-speech gestures might enhance memory more than arbitrary encodings in speech due to the visually motivated mappings between the linguistic form and the actual spatial configuration that it refers to (i.e., *iconicity*). In this study, we investigated the linguistic encoding and spatial memory accuracy of Left-Right spatial relations between objects, which are linguistically acquired late by speaking but not by signing children (Sümer, 2015). We asked whether linguistic encoding of the spatial relation between objects predicts better memory for the spatial relations for children. We also asked whether and how the modality of encoding (i.e., sign, speech, and speech-plus-gesture) modulates differences in spatial memory accuracy.

Children from early on start to communicate and reason about spatial relations (Landau, Dessalegn & Goldberg, 2011). Knowledge and use of spatial terms are strong predictors of children's spatial reasoning (Abarbanell & Li, 2021; Dessalegn & Landau, 2008; Gentner, et al., 2013; Loewenstein & Gentner, 2005; Hermer-Vasquez, Moffet & Muncholm, 2001; Miller, Andrews & Simmering, 2016; Shusterman, Lee & Spelke, 2011; Simms & Gentner, 2019). Using spatial terms such as *left-right* (Hermer-Vasquez, Moffet & Muncholm, 2001), *next* (Miller, Patterson, & Simmering, 2016), and *middle* (Simms & Gentner, 2019) helps children perform better in spatial memory tasks that contain these spatial relations. For instance, Lowenstein & Gentner (2005) found that children who received instructions containing spatial terms (i.e., *top, middle, bottom*) outperformed children who were not trained on these terms when tested for spatial memory both immediately and after a 2-day delay. However, this study lacks evidence as to whether the use of spatial terms by children themselves relates to their accuracy of spatial memory tasks that contain these spatial relations.

Moreover, this previous work investigating the relation between spatial language and spatial memory focused on the linguistic encoding of space in speech alone. However, language is a multimodal phenomenon and includes co-speech gestures and sign languages that operate on visual-spatial modality. Sign languages incorporate linguistic forms that bear visually motivated links to actual spatial relations (e.g., Emmorey, 2002 for American Sign Language and

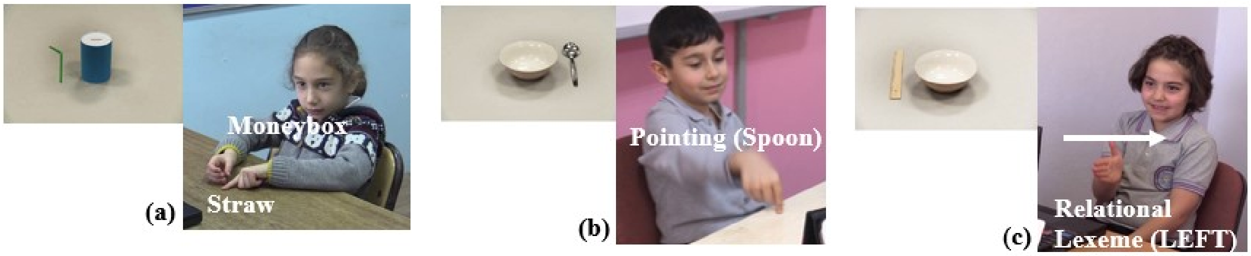
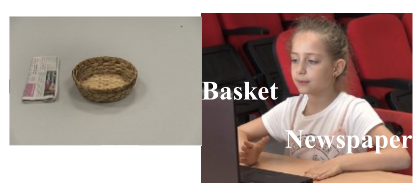


Figure 1. Examples from TİD signers encoding the spatial relation between (a) the moneybox and the straw using a classifier construction; (b) the bowl and the spoon by pointing to the location of the spoon on the signing space through an index finger; (c) the bowl and the ruler by using a relational lexeme for LEFT. *Note* that in all of these descriptions the participant introduced the lexical signs for the objects before depicting the spatial relation between the items via different linguistic forms.

Sümer, 2015 for Turkish Sign Language [TİD]). For instance, classifier constructions is one of the most frequent forms that signers use (Perniss, Özyürek & Zwitserlood, 2015) which employ handshapes resembling the size and shape of the objects and place these handshapes onto the signing space in an analog way to depict the spatial relation between the objects (Figure 1a). Signers may also use Other linguistic forms (Figure 1b&c) that encode spatial relations between objects on the signing space in iconic ways. However, these forms do not always encode information about size and shape of the objects involved (Karadöller, Sümer, Ünal & Özyürek, under review). Similarly, co-speech gestures can convey information regarding the spatial relations in an analog way together with speech (Figure 2). Despite the similarities between sign languages and co-speech gestures in describing spatial relations through visually motivated ways, they differ in the way they are acquired. Children acquiring sign languages are exposed to linguistic forms that bear visually motivated links as part of their language system. However, co-speech gestures are not learned as part of the linguistic system of spoken languages, rather they form a composite system together with speech (McNeill, 2005). Hence, there may be variation in the way children are exposed to gestures and the frequency and the type of gestures that are used by children. When it comes to speech, spatial relations between objects are represented through spatial terms that categorize spatial information in arbitrary ways. For instance, in order to describe the same picture in Figure 1a, an English speaker may say “the straw is to the *left* of the moneybox”. Therefore, speaking children need to learn arbitrary linguistic forms that encode space in categorical ways.

Until now, prior work did not take into account sign languages in investigating the relationship between spatial language and spatial reasoning of children. Moreover, it also underrepresented iconic forms in spoken languages by neglecting co-speech gestures (but see Abarbanell & Li, 2021 and Miller, Patterson & Simmering, 2016 that focused on gestures in isolation and compared them to speech directly). Focusing on speech-gesture combinations, however, is important to capture iconic aspects of spoken languages together with arbitrary speech forms when comparing spoken language descriptions to sign (Goldin-Meadow & Brentari, 2017; Özyürek & Woll, 2019).

Linguistic encoding of Left-Right is a late aspect of language development for speaking children (e.g., Benton, 1959; Harris, 1972; Piaget, 1972; Rigal, 1994; Sümer, 2015). However, recent research has shown that such delayed linguistic acquisition patterns obtained for speaking children are not present for children learning TİD (Sümer, Perniss, Zwitserlood & Özyürek, 2014). Earlier encoding of Left-Right spatial relations between objects by TİD-signing children compared to Turkish-speaking children has been attributed to the presence of visually motivated linguistic forms to encode spatial relations in sign (Sümer, 2015; Sümer et al., 2014). Furthermore, although Turkish-speaking 8-year-old children rarely use *left-right* spatial terms and instead use general spatial nouns (e.g., *side*) in their speech alone, they frequently use spatial co-speech gestures together with speech to encode spatial relation between the objects in iconic ways (see Figure 2; Karadöller, Sümer, Ünal & Özyürek, 2020). Therefore, iconic forms of communication might provide new evidence and unique insights into the relationship between spatial language and memory.



Sepet-in	yan-ın-da	gazete	var
Basket-GEN	side-POSS-LOC	newspaper	there-is

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

Figure 2. Example from a hearing child speaker of Turkish using spatial gestures representing the size/shape and the relative location of the objects while using less informative “side” in speech.

Even though not directly studied within the spatial domain, performing actions during encoding has been found to promote better memory representations than verbal encoding only for adults. This has been argued to be due to the involvement of the motor system leading to richer and/or stronger memory representations (see Cohen, 1989 and Nilsson, 2000 for reviews). Sign language descriptions and co-speech gestures may be comparable to performing actions since the execution of signs and co-speech gestures recruit hand movements that structurally resemble the movements executed in tasks that are used in performing actions. This claim has been supported by research showing that encoding through signing results in better memory accuracy compared to verbal encoding and this effect has been found to be as strong as the performing actions (von Essen & Nilsson, 2003; Zimmer & Engelkamp, 2003). Consequently, in addition to the iconicity of information conveyed in sign languages and co-speech gestures, the involvement of the motor system in executing these forms could lead to differences in spatial memory accuracy.

The Present Study

The present study investigates whether encoding a spatial relation in a picture description predicts later recognition memory accuracy for signing and speaking children. Besides, it investigates whether differences in the modality of encoding ("*sign vs speech alone*" or "*sign vs speech-plus-gesture*") lead to differences in recognition memory accuracy. To address these questions, hearing child speakers of Turkish and deaf child signers of TİD described pictures of objects in various spatial relations to each other, after which their memory for the described pictures was tested. We coded the picture descriptions for the presence of spatial relation in sign, speech, and co-speech gesture. We had the following predictions:

(a) Following previous research, we expect descriptions that convey spatial relation to have higher memory accuracy than descriptions that do not convey spatial relation regardless of the language of the participant.

(b) When the modality of encoding is concerned, we expect descriptions in sign to predict better memory accuracy than speech alone descriptions due to iconicity of the linguistic encoding and the activation of the motor system.

(c) When gestures are considered as part of the spoken descriptions, we foresee two possibilities. Descriptions in sign and speech-plus-gesture might generate equal memory accuracy. This might be because taking gestures into account might better represent spoken languages by incorporating iconic forms. Alternatively, signers might outperform speaking children even when speaking children's gestures are taken into account due to the conventional and obligatory ways to encode space in sign but not in co-speech gesture.

Method

The methods reported in this study have been approved by the Ethics Review Board of the Radboud University.

Participants Data were collected from 26 child monolingual speakers of Turkish (14 Female, *Mean Age* = 8;6) and 21 child deaf signers of TİD (12 Female, *Mean Age* = 8;5). Deaf children acquired their language from their deaf parents following birth. Participation was voluntary.

Materials Stimuli of the description task consisted of 84 displays with 4 pictures showing the same two objects (Figure and Ground) in various spatial configurations (i.e., Left, Right, Front, Behind, In, On). Ground objects (e.g., jar) was always in the center of the pictures and Figure objects (e.g., lemon) changed its location with respect to the Ground Objects. As a result, 4-pictures in one display differed only in terms of the position of the Figure object. This manipulation aimed to foreground spatial relations between the objects rather than the objects themselves and allowed us to ensure that participants mentioned the spatial relation between the objects as a distinguishing feature of the target picture in their descriptions. Each display had one target picture indicated by an arrow (Figure 3a).

Stimuli of the memory task consisted of a sample of the same displays (n = 54) without the arrow. Display order and the arrangement of the 4-pictures in one display were randomized for the recognition memory task (Figure 3b).

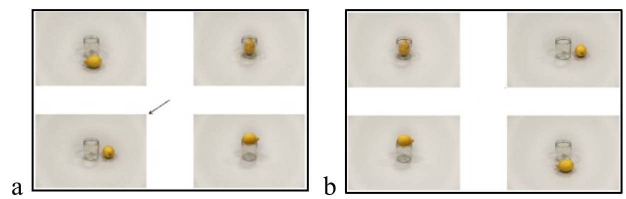


Figure 3. Example displays of the description (a) and memory (b) tasks.

Procedure Each participant was tested in a quiet room. All instructions were given orally in Turkish to speaking children by a Turkish speaking adult and in TİD to signing children by a deaf adult signer of TİD. No written instructions were used to avoid misunderstandings. The language status of the addressee was also matched with the participants'.

In the description task, each trial started with a fixation cross (2000ms), followed by a 4-picture display (1000ms). Next, an arrow was presented for 500ms targeting one of the pictures in the display (Figure 3a) and disappeared. Next, the display with 4-pictures remained on the screen for an additional 2000ms until the visual white noise screen appeared. Participants described the target picture to an addressee during the visual white noise. After the description, the addressee chose the picture that the participant described on her tablet. Participants moved to the next trial by pressing the ENTER key. This task was video recorded from side-top and front angles for later coding for the sign, speech, and gesture.

Following the description task, participants completed a Flanker Task as a distractor before the surprise recognition memory task. Adult participants received the original Flanker Task (duration 15mins) and child participants received the child-friendly version with fish (duration ~5mins).

In the surprise recognition memory task, participants were given the same displays and asked to click on the picture that they had previously described by using the mouse. The rationale for having this task as a surprise is to eliminate possible influences on the production performance. After the memory task, participants received the computerized version of the Corsi Block Tapping Task in forward order.

Coding All descriptions were transcribed for target pictures containing Left-Right spatial relation between the objects (n = 20). These descriptions later coded for the presence of spatial relation in sign, speech, and when speech and gestures were combined. Transcription and coding were done using ELAN (Sloetjes & Wittenburg, 2008).

In sign, participants used five different linguistic strategies to encode spatial relation between the objects such as classifier constructions (Figure 1a), pointing (Figure 1b), relational lexemes (Figure 1c), tracing of objects' size and shape on the signing space, and lexical verb placements). All of these forms provide the exact spatial information between the objects in iconic ways. In speech, participants used specific (e.g., *left*) or general spatial nouns (e.g., *side*). Participants also used spatial gestures together with their speech (Figure 2).

We grouped encodings by speakers into two categories based on the modality of information conveyed in the description to compare them with descriptions in sign. The first category (i.e., *speech alone*) incorporated the use of specific spatial nouns (i.e., *left-right*). Here, we excluded the descriptions with general spatial nouns (i.e., *side*) which did not specify the exact spatial relation between the objects. We made this selection to ensure that descriptions in sign and speech alone identified the pictures in similar ways and differed only in terms of the modality of encoding. The second category (i.e., *speech-plus-gesture*) consisted of speech gesture combinations that included specific or general spatial nouns with gestures.

Results

Control Measures

First of all, speaking and signing children were compared to ensure similarity in terms of age and visual-spatial working memory span (i.e., Corsi Block Tapping Task). Bayesian t-tests assessed the probability of the mean difference (MDIFF) greater than zero and less than zero using the R package BayesianFirstAid (version 0.1; Bååth, 2014). Signing and speaking children were similar in age (Bayesian two sample t-test: MDIFF(-5) > 0: $p = 0.556$, MDIFF(5) < 0: $p = 0.444$) and visual-spatial working memory (Bayesian two sample t-test: MDIFF(-5) > 0: $p = 0.972$, MDIFF(5) < 0: $p = 0.280$).

Memory Accuracy

Recognition memory data were analyzed with generalized binomial linear mixed effects modelling (*glmer*) with random intercepts for Subjects and Items. All models were fit using lme4 package (Bates, Mächler, Bolker & Walker, 2015) in R (R Core Team, 2018). This mixed effects approach allowed us to take into account the random variability that is due to having different participants and items. We did not include random slopes in our models because doing so either did not increase the model fit or the model was testing a between subject variable that cannot be added as a random slope.

Memory Accuracy in relation to Spatial Encoding Signing and speaking children varied in the frequency of encoding the spatial relations between items. Signing children encoded the spatial relation between items in 90% of their total descriptions via linguistic strategies available in sign languages. Speaking children encoded the spatial relation between the items in 96% of their total descriptions considering both speech (33%) and speech gesture combinations (63%).[1]

We analyzed whether having encoded the spatial relation between the items or not related to recognition memory accuracy of signers and speakers differently. We used a *glmer* model to test the fixed effects of Spatial Encoding (Spatial Encoding, No Spatial Encoding) and Language Group (Signer, Speaker) on binary values of Memory Accuracy (1 = Yes, 0 = No) at the item level (Figure 4). Fixed effects of Spatial Encoding and Language were analyzed with centered contrasts (-0.5, 0.5). The model revealed a fixed effect of Spatial Encoding ($\beta = 1.04$, $SE = 0.34$, $p < 0.01$) in which regardless of the Language Group, descriptions that encoded spatial relation related to higher memory accuracy compared to descriptions that did not contain a spatial encoding. There was no effect of Language Group ($\beta = 0.28$, $SE = 0.42$, $p = 0.52$) and no interaction between Spatial Encoding and Language Group ($\beta = 0.69$, $SE = 0.68$, $p = 0.31$).

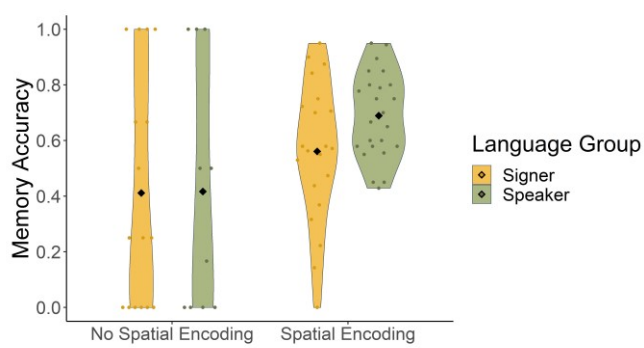


Figure 4. Memory accuracy of all descriptions across Language Group and Spatial Encoding.
Note. Here and the following graphs colored dots represent the mean memory accuracy for each participant. The black diamond represents mean memory accuracy of the group. The width of the violins represents the density of the data distribution. The length of the violins depicts the range of the data points.

[1] These include *left-right* spatial nouns (14%), general noun *side* (19%), and speech-plus-gesture combinations (63%).

Memory Accuracy in relation to Modality of Encoding Next, we investigated whether the modality of information conveyed in the descriptions of signers and speakers related to recognition memory accuracy. To do so, we did two comparisons: (1) Memory accuracy of descriptions with Sign vs Speech alone (2) Memory accuracy of descriptions with Sign vs Speech-plus-gesture.

First, we analyzed a subset of data which includes speech alone descriptions with *left-right* spatial nouns of speaking children (14% of the total descriptions) in comparison to all descriptions in sign for signing children (90% of the total descriptions). We tested to what extend the Modality of the encoding related to memory accuracy. We used a *glmer* model to test the fixed effect of Modality (Sign, Speech alone) on binary values of Memory Accuracy (1 = Yes, 0 = No) at the item level (Figure 5). The fixed effect of Modality was analyzed with centered contrasts (-0.5, 0.5). The model did not reveal a fixed effect of Modality ($\beta = 0.61$, $SE = 0.56$, $p = 0.29$), in which encoding via Sign or Speech alone did not lead to statistically significant differences in memory accuracy.

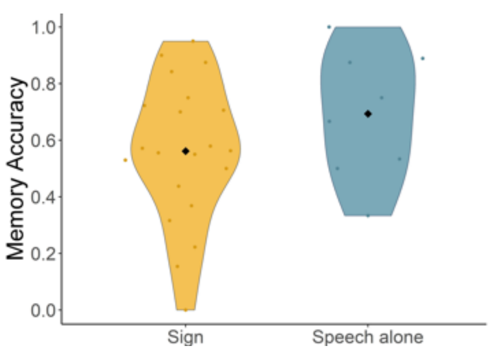


Figure 5. Memory accuracy of descriptions with Sign vs Speech alone.

Next, we analyzed speech-plus-gesture descriptions by speaking children (63% of the total descriptions) in comparison to all descriptions in sign for signing children (90% of the total descriptions). We tested to what extend the Modality of the encoding related to memory accuracy. We used a *glmer* model to test the fixed effect of Modality (Sign, Speech-plus-gesture) on binary values of Memory Accuracy (1 = Yes, 0 = No) at the item level (Figure 6). The fixed effect of the Modality was analyzed with centered contrasts (-0.5, 0,5). The model did not reveal a fixed effect of Modality ($\beta = 0.58$, $SE = 0.36$, $p = 0.11$) in which encoding via Sign or Speech-plus-gesture did not lead to statistically significant differences in memory accuracy.

These results suggest first of all that regardless of the language used in the description, encoding the spatial relation in the description related to higher recognition memory accuracy than not encoding a spatial relation. Moreover, the type of spatial information conveyed in the description via different modalities did not lead to statistically significant differences in memory accuracy.

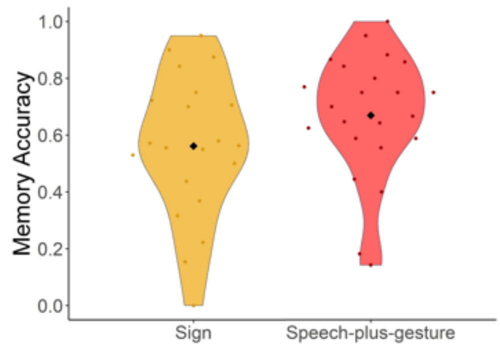


Figure 6. Memory accuracy of descriptions with Sign vs Speech-plus-gesture.

Discussion

In this study, we investigated whether linguistic encoding of the spatial relation between objects predicted better memory for the spatial relation between objects and whether this effect differs across signing and speaking children. Additionally, we tested to what extend the modality of encoding predicted memory accuracy. We had two key findings: First, encoding a spatial relation predicted higher memory accuracy both for signing and speaking children. Second, the modality of encoding (Sign vs Speech alone or Sign vs Speech-plus-gesture) did not predict statistically significant differences in memory accuracy.

First of all, as for the relation between spatial language and memory, our study is first to investigate the relation between producing spatial terms by children themselves to encode space and later accuracy in a memory task that involves these spatial relations. Unlike the previous work (e.g., Lowenstein & Gentner, 2005) that investigates training children in spatial terms and its relations to later memory accuracy, we show that spontaneous encoding of space by children themselves relates to better memory accuracy. With this finding, we also extend the literature on spatial language use and spatial memory abilities of children (e.g., Abarbanell & Li, 2021; Hermer-Vasquez, Moffet & Muncholm, 2001; Miller, Patterson & Simmering, 2016; Simms & Gentner, 2019) in a memory task that involves one-to-one correspondence to the linguistic encoding task. Moreover, our findings provide evidence to the view that encoding space through language may function as an additional medium of representation for space. This additional medium, in turn, result in enhanced representations and augment representational power (see Dessalegn & Landau, 2008; Gentner, 2016; see also Ünal & Papafragou, 2016 for a general discussion on language and cognition). Here, we showed that such an enhanced representation predicted stronger or richer memory as evidenced by higher accuracy when the spatial relation between the items were present regardless of the modality of the representation.

Our findings provide further evidence to the literature showing a strong relation between encoding with speech and later memory accuracy (Slamecka & Graf, 1978) and extend this to encoding through sign and gesture (see also Gentner et al, 2013 and ter Bekke, Özyürek & Ünal, 2019). We also extend these findings to the domain of space (Clark, 1973), where the modality of encoding might have an impact due to the visually motivated link between the linguistic form and its meaning while encoding via sign and gesture but not speech alone. However, it seems that regardless of the modality of encoding, once children encode the spatial relation in their descriptions, they have better memory representations compared to cases where they do not encode the spatial relation.

Despite the literature that shows a facilitating effect of encoding that incorporates actions on memory accuracy (Cohen, 1989; Nilsson, 2000) as well as the facilitating effect of encoding via sign over verbal encoding (Essen & Nilsson, 2003; Zimmer & Engelkamp, 2003), we did not find differences between memory accuracy of descriptions that are encoded via sign and speech alone. One possibility of such a lack of difference might result from the fact that speech alone descriptions consists of *left-right* spatial nouns. It is possible that despite the iconicity of encoding and the involvement of the motor system that might lead to stronger memory representations while describing via sign, encoding the Left-Right spatial relations between objects by using spatial terms have their own challenges for speaking children. It might be that encoding space through spatial terms that categorize the spatial relation in arbitrary ways requires more effortful processing than simply depicting the spatial relation in iconic ways. Hence, results in memory representations that are equally strong to encoding via iconic forms. This is especially plausible for descriptions of Left-Right relations in which children reported to have difficulties in cognitive and linguistic encoding (Benton, 1959; Harris, 1972; Piaget, 1972; Sümer, 2015) possibly due to the lack of any pragmatical cues to differentiate between symmetrical nature of Left and Right (see Grigoroglou, Johanson, & Papafragou, 2019 for the discussion on pragmatical cues to distinguish Front and Back).

Moreover, similar memory accuracy of descriptions in sign and speech alone might have resulted from the experimental differences of the verbal encoding condition as well as the type of stimuli items used to generate encodings. Previous studies showing a facilitating effect of encoding in sign over verbal encoding considered reading silently (Zimmer & Engelkamp, 2003) and hearing lists of words that were read to participants (von Essen & Nilsson, 2003) as the verbal encoding, in which participants were not actively producing stimuli items as our participants did. It has been found that having produced spoken descriptions generated better memory than words that were read silently (i.e., the production effect; e.g., Conway & Gathercole, 1987). Hence, in our study, having produced descriptions aloud might have generated stronger memory representations that can compete with encoding in sign. Besides, in contrast to previous studies that included lists of words to be encoded, our stimuli consisted of pictures of items. It has been robustly reported that pictures generates better memory than lists of words (i.e., the picture superiority effect; e.g., Paivio, Rogers & Smythe, 1968). Possibly, using pictures as stimuli items induced more unique representations than words (Nelson, Reed & McEvoy, 1977) or generated extensive conceptual processing of the stimuli compared to words (Weldon & Roediger, 1987).

Finally, encoding via sign and speech-plus-gesture also predicted equal memory accuracy. It seems that taking gestures into account together with speech represented spoken languages on equal grounds to sign languages by not only focusing on arbitrary speech forms but also incorporating iconic forms of communication represented by co-speech gestures (Goldin-Meadow & Brentari, 2017; Özyürek & Woll, 2019).

In conclusion, the present study contributes new evidence to the link between the use of spatial language and later memory accuracy of children. Our findings show for the first time that the use of spatial language by children themselves predicts better accuracy in memory tasks that require these spatial relations. However, the modality of encoding via sign, speech, or speech-plus-gesture does not further modulate differences in memory accuracy of children.

Acknowledgements

This work has been supported by NWO VICI Grant awarded to the last author. We thank our deaf assistants Sevinç Yücealtay-Akın, Feride Öksüz, Yusuf Ermez and our hearing assistants Hükümran Sümer, İrem Yeşilçavdar, Yeşim Özüer for their help with data collection, annotation, and coding.

References

Abarbanell, L., & Li, P. (2021). Unraveling the contribution of left-right language on spatial perspective taking. *Spatial Cognition & Computation*, 1-38.

Bååth, R. (2014). BayesianFirstAid: A package that implements bayesian alternatives to the classical test functions in R.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Benton, A. (1959). *Right-left discrimination and finger localization.* New York: Hoeber-Harper.

Clark, E. V. (1973). Non-linguistic strategies and the acquisition of word meanings. *Cognition, 2*(2), 161-182.

Cohen, R. L. (1989). Memory for action events: The power of enactment. *Educational psychology review, 1*(1), 57-80.

Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. J*ournal of Memory and Language, 26*, 341–361.

Dessalegn, B., & Landau, B. (2008). More than meets the eye: The role of language in binding and maintaining feature conjunctions. *Psychological science, 19*(2), 189-195.

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

Gentner, D. (2016). Language as Cognitive Tool Kit: How Language Supports Relational Thought. *American Psychologist, 71*(8), 650–657.

Gentner, D., Özyürek, A., Gürcanli, Ö., & Goldin-Meadow, S. (2013). Spatial language facilitates spatial cognition: Evidence from children who lack language input. *Cognition, 127*(3), 318-330.

Grigoroglou, M., Johanson, M., & Papafragou, A. (2019). Pragmatics and spatial language: The acquisition of front and back. *Developmental Psychology, 55*(4), 729.

Goldin-Meadow, S., & Brentari, D. (2017). Gesture, sign, and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences, 40*, e46.

Harris, L. (1972). Discrimination of left and right, and development of the logic relations. *Merrill-Palmer Quarterly, 18,* 307-320.

Hermer-Vazquez, L., Moffet, A., & Munkholm, P. (2001). Language, space, and the development of cognitive flexibility in humans: The case of two spatial memory tasks. *Cognition, 79*(3), 263–299.

Karadöller, D. Z., Sümer, B., Ünal, E., & Özyürek, A. (Under Review). Late sign language exposure does not modulate the relation between spatial language and spatial memory in deaf children and adults.

Karadöller, D. Z., Sümer, B., Ünal, E., & Özyürek, A. (2020). Sign advantage for children: Signing children's spatial expressions are more informative than speaking children's speech and gestures combined. The 45th Annual Boston University Conference on Language Development, Boston, USA.

Landau, B., Dessalegn, B., & Goldberg, A. M. (2011). Language and Space: Momentary Interactions. in P. Chilton and V. Evans (Eds.), *Language, cognition and space: The state of the art and new directions. Advances in Cognitive Linguistics Series.* London: Equinox Publishing.

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive psychology, 50*(4), 315-353.

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

Miller, H. E., Patterson, R., & Simmering, V. R. (2016). Language supports young children's use of spatial relations to remember locations. *Cognition*, 150, 170-180.

Nelson, D. L., Reed, V. S., & McEvoy, C. L. (1977). Learning to order pic-tures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning and Memory, 3*(5), 485–497.

Nilsson, L. G. (2000). *Remembering actions and words.* The Oxford handbook of memory.

Özyurek, A., & Woll, B. (2019). Language in the visual modality: Cospeech gesture and sign language. In P. Hagoort (Ed.), *Human language: From genes and brain to behavior*. Cambridge, MA: MIT Press.

Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures to recall than words? *Psychonomic Science, 11*(4), 137–138.

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

Piaget, J. (1972). *Judgment and reasoning in the child.* Totowa, NJ: Littlefield, Adams. (Originally published in 1928).

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rigal, R. (1994). Right-left orientation: Development of correct use of right and left terms. Perceptual and Motor Skills, 79(3), 1259.

Shusterman, A., Lee, S. A., & Spelke, E. S. (2011). Cognitive effects of language on human navigation. *Cognition, 120*(2), 186-201.

Simms, N. K., & Gentner, D. (2019). Finding the middle: Spatial language and spatial reasoning. *Cognitive Development, 50*, 177–194.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of experimental Psychology: Human learning and Memory*, *4*(6), 592-604.

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category - ELAN and ISO DCR. *Proceedings of the 6th International Conference on Language Resources and Evaluation* (LREC 2008).

Sümer (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, Nijmegen.

Sümer, B., Perniss, P.M., Zwitserlood, I.E.P., & Özyürek, A. (2014). Learning to express "left-right" & "front-behind" in a sign versus spoken language. In P. Bello, M. Guarini, M. McShane & B. Scassellati (Eds.), *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*. Austin,Tx: Cognitive Science Society.

ter Bekke, M., Ozyurek, A., & Ünal, E. (2019). Speaking but not gesturing predicts motion event memory within and across languages. In A. Goel, C. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci 2019)* (pp. 2940-2946). Montreal, QB: Cognitive Science Society.

Ünal, E., & Papafragou, A. (2016). Interactions between language and mental representations. *Language Learning*, 66(3), 554–580.

von Essen, J. D., & Nilsson, L. G. (2003). Memory effects of motor activation in subject-performed tasks and sign language. *Psychonomic bulletin & review, 10*(2), 445-449.

Weldon, M. S., & Roediger, H. L. 3rd. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory & Cognition, 15*(4), 269–280.

Zimmer, H. D., & Engelkamp, J. (2003). Signing enhances memory like performing actions. *Psychonomic Bulletin & Review, 10*(2), 450-454.