

# Speaker-Related Cognitive Constraints on Multimodal Audience Design During Spatial Communication

**Demet Özer (demet.ozer@bilkent.edu.tr)**

Department of Psychology, Bilkent University  
Universiteler Çankaya, 06800, Ankara - Türkiye

**Tilbe Göksun (tgöksun@ku.edu.tr)**

Department of Psychology, Koç University  
Rumelifeneri Yolu Sarıyer, 34450, Istanbul - Türkiye

## Abstract

Human communication is multimodal, characterized by the use of speech and co-speech iconic gestures. While previous research examined how cognitive and communicative demands affect multimodal language use, the interplay between speaker-related cognitive constraints and listener-oriented adaptations remains unclear. The present study examines how speakers with varying spatial skills adjust their speech and iconic gestures when addressing interlocutors with different spatial skills. By employing the imagined addressee paradigm, twenty-three participants described how to solve mental rotation problems to interlocutors with low- versus high-spatial skills. Speakers produced more iconic gestures when addressing low-spatial compared to high-spatial interlocutors. However, this adaptation was primarily observed among individuals with above-average spatial skills. These findings extend prior work on multimodal audience design, showing that while speakers adjust their gestures based on listener characteristics, these adaptations are constrained by the speaker's cognitive capacities. This study highlights the dynamic interaction between cognitive and communicative demands in multimodal language.

**Keywords:** multimodal communication, gesture, audience design, spatial skills

## Introduction

Face-to-face human communication is inherently multimodal, characterized by the simultaneous use of multiple communicative signals. These signals stem from different modalities such as speech and co-speech iconic hand gestures that illustrate the physical properties of objects, actions, and events (Holler & Levinson, 2019; McNeill, 2005; Vigliocco et al., 2014). Iconic gestures serve many functions during language production and comprehension, information encoding, learning, and problem-solving processes. Functionalist theories that seek to explain why speakers use gestures offer two broadly categorized functions for gestures: speaker-oriented cognitive functions and listener-oriented communicative functions (Goldin-Meadow et al., 2001; Kita & Özyürek, 2003; Pouw et al., 2014; Cook & Fenn, 2017; Kita et al., 2017; Novack & Goldin-Meadow, 2017). Speaker-oriented theories assert that gestures primarily serve speakers' own cognitive processes and facilitate thinking and speaking. Gestures help activate,

maintain, and package visual, spatial, and motoric information (Kita et al., 2017). Moreover, gestures reduce cognitive load by keeping information active in the working memory and projecting internal representations to the external physical space (Marstaller & Burianova, 2013; Pouw et al., 2014). Given this, speakers may use gestures as a compensatory and supporting tool to manage cognitive load or limited cognitive resources (e.g., lower cognitive skills, Chu & Kita, 2011; Özer & Göksun, 2020a). Listener-oriented theories, on the other hand, claim that gestures primarily serve communicative functions and speakers may use gestures to convey meaning tailored for their interlocutors (Goldin-Meadow & Alibali, 2013). Speakers adjust their gestures depending on their listeners' global characteristics and communicative needs, a principle known as Audience Design (Clark & Murphy, 1982; Holler & Stevens, 2007). Speaker- and listener-oriented functions of gestures are not mutually exclusive. Rather, gestures serve both functions simultaneously. Considering these, two key factors may influence gesture use during communication: (1) speakers' cognitive skills and (2) listener-related factors such as global characteristics, needs, and perceived cognitive profiles of interlocutors.

Although earlier research present evidence for the role of each factor in isolation (e.g., Campisi & Özyürek, 2013; Chu et al., 2014; Göksun et al., 2013; Kandemir et al., 2023; Özer & Göksun, 2020a), the dynamic interplay between listener-related factors and speaker-related cognitive constraints is mostly unknown. How do speakers with varying cognitive skills adapt their gestures when interacting with interlocutors of varying cognitive skills? Addressing this question might yield a more comprehensive account for understanding factors underlying multimodal language production, where both cognitive and social-communicative demands are at play. To fill this gap, the present study employs imagined addressee paradigm and examines the interplay among speaker- and listener-related factors on multimodal language in a spatial context, where gestures are particularly prevalent and effective (Alibali, 2005). We investigate whether and how (1) speakers adjust their speech and co-speech iconic gestures when addressing interlocutors with varying spatial skills and (2) the role of speakers' spatial skills in this process

during spatial descriptions (i.e., explanations of mental rotation problems).

### **Listener-Oriented Adaptations in Multimodal Communication**

Speakers adapt their language to the needs, beliefs, and knowledge of their interlocutors (Clark & Murphy, 1982). For instance, speakers tend to produce longer, more informative, and structurally more complex utterances when there is no common ground between interlocutors (Blokpoel et al., 2012; Ferreira, 2019; Galati & Brennan, 2010). Similarly, adult speakers tend to use shorter sentences and simplified vocabularies when talking to a child compared to an adult (Brand et al., 2002; Iverson et al., 1999). This design principle is not limited to speech but also plays a prominent role in the use of other communicative signals, such as iconic gestures. Recent studies showed that speakers modify not only their speech but also their co-speech iconic gestures depending on the global characteristics and needs of their listeners (Campisi & Özyürek, 2013; Gerwing & Bavelas, 2004; Hoetjes et al., 2015; Holler & Stevens, 2007; Holler & Wilkin, 2009; Jacobs & Garnham, 2007; Kandemir et al., 2023; Schubotz et al., 2019; Tellier et al., 2021). For example, speakers use more iconic gestures when narrating a story to a listener who has not heard the story before compared to one who has heard it before (Holler & Stevens, 2007; Holler & Wilkins, 2009; Jacobs & Garnham, 2007). These gestures are not only more frequent but semantically more complex (Gerwing & Bavelas, 2004). Studies that employed the imagined-addressee paradigm showed increased use of iconic gestures by adult speakers when addressing to a child compared to an adult (Campisi & Özyürek, 2013; Kandemir et al., 2023). In a similar vein, speakers tend to use more iconic gestures that are produced for longer durations and in larger gestural spaces when addressing non-native listeners, who are assumed to have lower verbal proficiency, compared to native listeners (Tellier et al., 2021; Prove et al., 2022). Taken together, these suggest that speakers strategically modify their speech and co-speech iconic gestures depending on the global characteristics and needs of their interlocutors. A key open question, however, is whether and how speaker-related cognitive factors interact with listener-oriented adaptations. The current study aims to examine the interplay between speakers' and their imagined addressees' perceived cognitive skills during multimodal language production.

### **Speaker-Related Cognitive Constraints on Multimodal Communication**

Drawing from speaker-oriented functionalist theories that assert the enhancing role of gestures in speakers' thinking and speaking processes, numerous studies have shown that individuals with lower cognitive abilities tend to use more iconic gestures during language production and problem solving and benefit more from using these gestures in learning processes (Chu et al., 2014; Gillespie et al., 2014;

Marstaller & Burianova, 2013; see Özer & Göksun, 2020a for a review). One important cognitive resource that contributes to variation across individuals in the production and comprehension of gestures is spatial skills (e.g., spatial working memory, spatial mental imagery, mental rotation ability; Abramov et al., 2021; Chu et al., 2014; Göksun et al., 2013; Hostetter & Alibali, 2007; Özer & Göksun, 2020b; Özer et al., 2025; Trafton et al., 2006). Research has shown that individual differences in spatial skills are linked to gesture use. For example, individuals with lower spatial working memory capacity, mental rotation ability, and spatial conceptualization ability used more gestures when explaining abstract concepts or social dilemmas compared to individuals with higher spatial skills (Chu et al., 2014). Similarly, when asked to describe how to solve mental rotation problems, individuals with lower spatial skills (i.e., lower mental rotation performance) used more gestures than those with higher spatial skills (Göksun et al., 2013). Additionally, the types of gestures used by low and high spatial ability speakers differed. Low spatial ability speakers used more static gestures referring to the entire object (e.g., cube, whole figure) when explaining mental rotation problems, whereas high spatial ability speakers used more dynamic gestures indicating movement and direction (e.g., moving to the right) as well as gestures referring to parts of the figure (e.g., the L-shaped part). Importantly, gesture use has also been evidenced to benefit spatial problem solving. All these studies demonstrate a link between spatial skills and iconic gesture use, particularly in contexts rich in spatial information. Chu and Kita (2011) showed an increased mental rotation performance when participants were encouraged to use gestures during problem solving compared to when they used gestures spontaneously.

Considering the supportive role of gestures in speech and thought processes, individuals with lower cognitive skills particularly benefit more from this reinforcing effect of gestures and use gestures as a tool to reinforce their cognitive processes during communication and problem solving (Marstaller & Burianova, 2013; Özer & Göksun, 2020a; Pouw et al., 2014).

Although there is no direct evidence on the interaction between speakers' and listeners' cognitive characteristics in determining multimodal language use, earlier studies suggest that speakers' cognitive skills may play a role in listener-oriented communication design processes. For example, young and elderly adults were asked to recount a story to a listener whom they believed had either heard it before or was hearing it for the first time (Horton & Spieler, 2007; Schubotz et al., 2019). Young adults used fewer words and gestures when they believed the listener had already heard the story compared to when they believed the listener was hearing it for the first time. Elderly adults, however, used comparable number of words and gestures to both listeners. Due to age-related declines in verbal and spatial abilities (Andersen & Ni, 2008; Copeland & Radvansky, 2007), elderly adults may use gestures for their own speaking and thinking processes, making them less likely to adjust their multimodal language

use based on shared knowledge with the listener (Göksun et al., 2022; Schubotz et al., 2019). Similarly, Kandemir and colleagues (2023) showed that adult speakers used more iconic gestures when recounting a story to a child compared to an adult, whereas they used comparable number of iconic gestures to a child and an adult in a map task, where they described how to navigate from one location to another on a map. This finding suggest that in contexts where gestures are particularly effective (i.e., route descriptions, Alibali, 2005), speakers may use gestures for their own cognitive processes, leading to diminished listener-oriented adaptations in such cases.

## The Present Study

The present study examines the interplay between speaker- and listener-related factors on multimodal language use (i.e., speech and co-speech iconic gesture use) during explanations in a spatial task. Specifically, we investigate whether and how (1) speakers adjust their speech and iconic gestures depending on the perceived spatial aptitude of their interlocutors (i.e., when addressing to an imagined interlocutor with low versus high mental rotation performance) and (2) speakers' spatial skills interact with the listener-related adaptations in multimodal language production. To address these questions, we conducted a controlled experiment where participants first completed the computerized mental rotation task to assess their spatial skills. They were then asked to describe how to solve some different mental problems to two imagined addressees with varying mental rotation problems: low- versus high-spatial skill audience. We coded the spatial terms in speech, as well as iconic and beat gestures. For the research questions posed above, we hypothesized that (1) participants would use more spatial speech and iconic gestures when addressing the low-spatial-audience compared to high-spatial-audience, and (2) if gestures serve cognitive functions for speakers and individuals with lower spatial skills benefit from using gestures for their own cognitive processes during speaking and thinking, listener-related adaptation would be present only for speakers with higher skills. That is, individuals with higher spatial skills would use more iconic gestures when addressing the low- than high-spatial audience, whereas individuals with lower spatial skills would use comparable number of iconic gestures across both audience.

## Methods

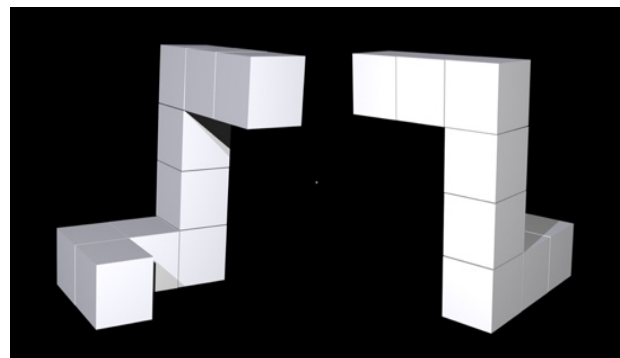
### Participants

Twenty-four native-Turkish speaking adults (16 females,  $Mage = 20.7$ ,  $SD = 3.6$ , range = 18-29) participated in the study in exchange for course credit. The sample size was based on power analyses conducted on G\*Power 3.1 (Faul et al., 2007), with an effect size of .386 (computed based on partial eta square of .13 for iconic gesture use across child-versus adult-directed speech in Kandemir et al., 2023) and a power of .95 for repeated measures ANOVA with 2 (audience order: low-audience-first versus -second) \* 2

(audience type: low-spatial-audience versus high-spatial-audience), which yielded 24 participants. All participants were right-handed and had normal or corrected-to-normal vision. Before the experiment, they gave written informed consent in accordance with the policies of Koç University's Institutional Ethics Committee on Human Research. One of the participants was discarded for having extremely low mental rotation task score (lower than two standard deviations below the group mean). Thus, the final sample consisted of 23 participants (15 females,  $Mage = 20.7$ ,  $SD = 3.7$ ).

### Tasks and Stimuli

**Mental Rotation Task** We used the computerized version of mental rotation task (Shepard & Metzler, 1971) to assess participants' spatial skills. In this task, participants were shown two three-dimensional figures that were rotated by different angles (0, 50, 100, or 150 degrees) side by side on a computer screen (Figure 1). Participants were asked to mentally rotate these figures and decide whether they were the same or different by pressing one of two keys on the keyboard (N for "same," V for "different"). Participants completed a total of 96 trials, with 12 unique figures each rotation angles (4 categories: 0,50,100, or 150) and response type (2 categories: same or different). All trials were presented in a randomized order. Participants completed 15 practice trials to ensure they understood the task and were given feedback on their responses during this practice phase.



**Figure 1:** Sample trial for the mental rotation task (50 degrees)

**Explanation Task for Mental Rotation Questions** After the completion of the computerized mental rotation task, participants were asked to explain how to solve some mental rotation problems similar to the ones described above to two different addressees, both of whom had previously completed the computerized mental rotation task. Participants' explanations were video recorded. They were informed that their explanations would be shown to their two addressees who would use these explanations to help them solve the questions later. One listener was presented as having relatively low mental rotation performance (52% accuracy),

while the other was shown as having relatively high performance (96% accuracy). The terms “low” or “high” performance were not explicitly mentioned and only their performance in terms of accuracy were presented together to facilitate comparison. Each participant explained six different questions (2 per 50, 100, and 150 degrees) for each addressee. The order of the addressee was counterbalanced across participants. Half of the participants explained to the low-spatial addressee first, while the other half explained to the high-spatial addressee first. After completing the first set of questions to the first addressee, participants were reminded of the second addressees’ performance and asked to explain a different set of questions to this listener. The question sets were counterbalanced across addressee types and questions within each set were randomly presented.

## Procedure

After completing the demographic form, participants first completed the computerized mental rotation task, which was conducted using E-Prime software. Then they completed the explanation task, which was presented using Microsoft PowerPoint. Participants sat on an armless and stationary chair and were asked to describe how to solve mental rotation problems spontaneously by directly looking at the camera. They were not given any explicit instruction on gesture use not to confound spontaneous use of gestures. The entire session was video recorded and lasted approximately 30 minutes.

## Transcription and Coding

**Speech** Participants’ speech was transcribed verbatim using the ELAN software. We calculated the total number of words uttered for each trial. Next, we coded spatial speech: (1) static spatial terms that describe within-object spatial properties (e.g., U-shaped figure, corner), (2) location and direction terms that describe movement directions (e.g., rightward) or relative spatial locations (e.g., the bottom), and (3) spatial knowledge terms that refer to spatial attributes (e.g., horizontal axis, symmetry). We computed a composite spatial speech score by summing all these categories and then calculating the total number of spatial terms per 10 words uttered for each trial.

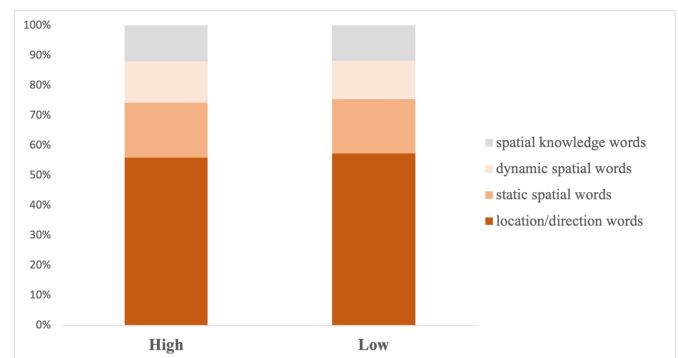
**Gesture** We coded participants’ spontaneous iconic and beat gestures. Iconic gestures represented the figure itself (e.g., showing an L-shaped figure with the hand), rotation of the figure (e.g., showing an L-shaped figure while rotating it rightward), or the location of the figure (e.g., holding the hand to the upper right hand to show the location of the cube). There were also pointing gestures that indexically showed a certain location, such as pointing to the right-hand side to express the rotation direction. These pointing gestures were considered as part of iconic gestures. Beat gestures were rhythmical hand movements that carried no meaning. We computed the total number of iconic and beat gestures used per 10 words uttered for each trial.

**Reliability** There were two independent coders in the study. The first trained coder coded all the data whereas the second one coded only 20% of the data (5 participants) to obtain inter-rater reliability. For iconic gesture, the Intraclass Correlation Coefficient (ICC) between the coders was very high,  $ICC = 0.94$ ,  $CI = [0.67, 0.99]$ ,  $p < .001$ . For spatial speech, there was also high reliability between the coders,  $ICC = 0.98$ ,  $CI = [0.90, 1.00]$ ,  $p < .001$ .

## Results

We calculated the number of iconic gestures, beat gestures, and spatial speech used per 10 words uttered in total and conducted three separate linear-mixed effects models for each as an outcome variable. The fixed effects included the audience order (low-audience-first vs. -second), audience (high- vs. low-audience), mental rotation scores, and the two-way interactions between the audience order and audience, and the audience and mental rotation scores. The random effects included random intercepts for items and participants.

**Spatial Speech Use.** There was only a significant interaction between audience and audience order in the composite spatial speech use,  $F(1, 249.02) = 4.92$ ,  $p = .027$ . However, Bonferroni-corrected post hoc analyses showed no differences across conditions. The frequencies of specific spatial speech categories were comparable across audience types, all  $ps > .05$ . Across both audience types, speakers used location and direction spatial terms the most, followed by static spatial terms (see Figure 2).

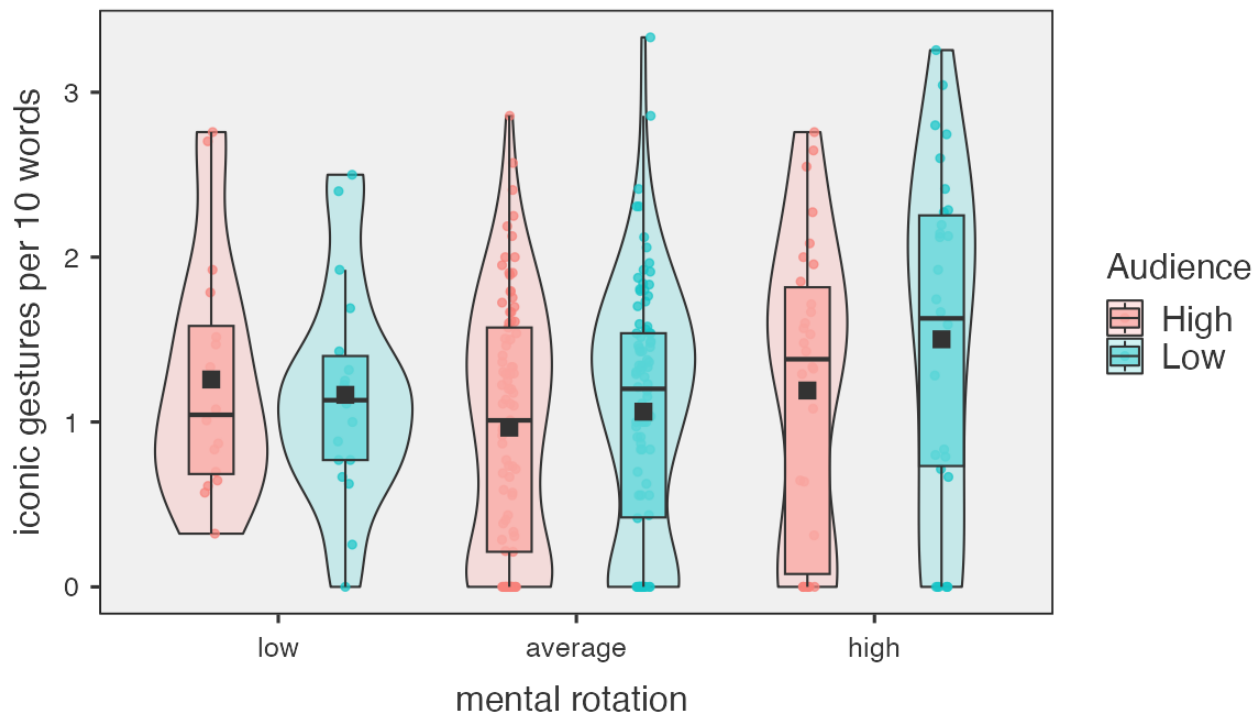


**Figure 2:** Distribution of specific spatial speech categories across low- and high-spatial- audience

**Iconic gesture use.** Fixed effects omnibus tests revealed a significant main effect of the audience,  $F(1, 242) = 4.79$ ,  $p = .030$ . Participants used more iconic gestures to low-audience compared to high-audience by  $1.73 \pm 0.79$ ,  $t(242) = 2.19$ ,  $p = .030$ . There were no main effects of the audience order ( $F(1, 20) = 0.13$ ,  $p = .727$ ) and mental rotation scores,  $F(1, 20) = 0.06$ ,  $p = .803$ . There was a significant interaction between the audience and mental rotation scores,  $F(1, 242) = 5.24$ ,  $p = .023$ . Simple effects analyses showed that the iconic gesture use across low- vs. high-

audience differed ( $M_{diff} = 0.23, SE = 0.10$ ) only among individuals with above-average (mean + 1 SD) mental rotation scores,  $F(1, 242) = 5.65, p = .018$ . Iconic gesture use did not differ across low- vs. high-audiences among participants with below-average ( $F(1, 239) = 0.76, p = .384$ ) and average mental rotation performance,  $F(1, 238) = 1.13, p = .288$ . There was also a significant interaction between the audience order and audience,  $F(1, 238) = 5.60, p = .019$ . However, Bonferroni-corrected post-hoc tests revealed no differences across conditions, all  $p$ s > .082. Figure 3 presents iconic gesture use for low- vs. high-spatial-audience across participants with below-average, average, and above-average mental rotation skills.

**Beat gesture use.** There was a main effect of the audience order,  $F(1, 19.99) = 4.52, p = .046$ . Participants used more beat gestures in the low-audience-first compared to high-audience-first condition by  $0.25 \pm 0.12, t(19.99) = 2.12, p = .046$ . There were no main effects of the audience and the mental rotation scores,  $F(1, 242.95) = 2.12, p = .147$  and  $F(1, 19.98) = 0.12, p = .732$ , respectively. The interactions between audience order and audience, and between audience and mental rotation scores were not significant as well,  $F(1, 238.93) = 0.82, p = .367$  and  $F(1, 243.23) = 2.05, p = .154$ , respectively.



**Figure 3:** The number of iconic gestures (per 10 words) used when interacting with low-spatial-audience versus high-spatial-audience across participants who had low (mean – 1SD), average (mean), and high (mean + 1SD) mental rotation performance. Each dot represents a single trial, illustrating 276 data points in total (23 participants X 2 audience X 6 trials per each audience).

### Discussion

The present study examined how speakers adjusted their multimodal language, specifically speech and co-speech iconic gestures, when addressing imagined interlocutors with varying spatial skills. More specifically, we asked whether and how speakers modified their multimodal language based on their imagined addressees' perceived spatial skills when explaining mental rotation problems and whether speakers' own spatial skills interacted with this process. By doing so, we aimed to examine the dynamic interplay between listener-related communication factors and speaker-related cognitive constraints during spatial multimodal language use. Our

results revealed that speakers used more iconic gestures when addressing a low-spatial audience compared to a high-spatial audience. However, this adaptation was primarily observed among individuals with above-average spatial skills. In contrast, those with lower spatial skills used comparable numbers of iconic gestures across both audience types. This adaptation was not present in spatial speech and beat gesture use, evident in comparable frequencies across both audience types.

Our findings provide corroborating evidence for earlier studies suggesting that audience design principle is not only present in speech but also extends to other communicative cues, particularly iconic gestures (Campisi & Özyürek, 2013;

Gerwing & Bavelas, 2004; Holler & Stevens, 2007; Holler & Wilkins, 2009; Kandemir et al., 2023). Although earlier evidence showed multimodal adjustments depending on the global characteristics (e.g., age such as interacting with a child versus adult; Campisi & Özyürek, 2013; Kandemir et al., 2023) and informational needs (e.g., common ground between interlocutors, Holler & Stevens, 2007; Holler & Wilkins, 2009) of listeners, the present study, for the first time, present evidence for gestural adaptation depending on a rather specific cognitive profile, namely perceived spatial aptitudes of interlocutors in a mental rotation task. Moreover, our findings showed that this adaptation primarily occurs for speakers with higher spatial skills. This suggests that speakers with higher spatial skills can adjust their gestures more flexibly to accommodate the communicational needs of their audience, whereas speakers with lower spatial skills did not exhibit this flexibility. This aligns with previous work suggesting that individuals with lower cognitive resources may rely on gestures primarily for their own cognitive processes (Marstaller & Burianova, 2013; Özer & Göksun, 2020a; Pouw et al., 2014). That is, while gestures can serve communicative functions, their use may be partially limited by the speaker's cognitive demands, limiting the extent of the listener-oriented adaptations. Overall, our results present key contributions to the field by suggesting that whether and how speakers adjust their gestures for their interlocutors can be constrained by the speaker's cognitive demands.

While speakers modified their iconic gestures depending on their imagined addressees' perceived spatial skills, their beat gestures and spatial speech remained stable based on listener characteristics for both low- and high-spatial skilled individuals. The lack of findings in the beat gesture use suggests that audience design is not a general feature of multimodal communication but rather specific to cues that carry semantic information. Unlike iconic gestures that convey semantic meaning through their form, beat gestures are primarily linked to speech fluency and prosodic organization (Ferrari & Hagoort, 2025; Kraemer & Swerts, 2007).

Despite these promising findings, some limitations should be acknowledged. First, the current study has a relatively small sample size, which may limit the generalizability of the findings. However, the current study presents preliminary data from an ongoing project and we aim to present data from a larger sample size in our future work. Additionally, the current study employed the imagined addressee paradigm, in which participants were asked to address an imagined interlocutor. Although this paradigm has been extensively used in the literature (e.g., Campisi & Özyürek, 2013; Kandemir et al., 2023) and provides control for local feedback (e.g., online feedback from an actual interlocutor such as head nods) that would otherwise present potential confounds, real-time interactions might yield different patterns for multimodal audience design. Future studies could extend this work by examining interactive communicational settings.

In conclusion, the current research underscores the importance of considering both speaker-related and listener-related factors in multimodal language production. Although speakers adjust their multimodal language depending on listeners' cognitive skills, these adaptations are not monolithic and can be constrained by speakers' own cognitive skills. These findings highlight the dynamic and complex interplay between cognition and communication.

## Acknowledgements

We thank Arda İçen, Zeynep Kaya, and Berkcan Işık for their help in coding transcription and coding, and Fatma Elif Müjdecı for data preparation. We also thank the members of Multimodal Language Lab at Bilkent University and Language & Cognition Lab at Koç University for their helpful feedback.

## References

- Abramov, O., Kern, F., Koutalidis, S., Mertens, U., Rohlfing, K., Kopp, S. (2021). The relation between cognitive abilities and the distribution of semantic features across speech and gesture in 4-year-olds. *Cognitive Science*, 45(7), e13012.
- Alibali, M. W. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation*, 5(4), 307-331.
- Andersen, G. J., Ni, R. (2008). Aging and visual processing: Declines in spatial not temporal integration. *Vision Research*, 48(1), 109-118.
- Blokpoel, M., van Kesteren, M., Stolk, A., Haselager, P., Toni, I., Van Rooij, I. (2012). Recipient design in human communication: simple heuristics or perspective taking. *Frontiers in Human Neuroscience*, 6, 253.
- Brand, R. J., Baldwin, D. A., Ashburn, L. A. (2002). Evidence for 'motionese': modifications in mothers' infant-directed action. *Developmental Science*, 5(1), 72-83.
- Campisi, E., Özyürek, A. (2013). Iconicity as a communicative strategy: Recipient design in multimodal demonstrations for adults and children. *Journal of Pragmatics*, 47(1), 14-27
- Chu, M., Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General*, 140(1), 102.
- Chu, M., Meyer, A., Foulkes, L., Kita, S. (2014). Individual differences in frequency and saliency of speech-accompanying gestures: the role of cognitive abilities and empathy. *Journal of Experimental Psychology: General*, 143, 694-709.
- Clark, H. H., Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in Psychology*, 9, 287-299.
- Cook, S. W., Fenn, K. M. (2017). The function of gesture in learning and memory. In Church, R.B., Alibali, M.W. & Kelly, S.D (Eds.), *Why gesture? How the hands function in*

- speaking, thinking, and communicating*. Amsterdam: John Benjamins Publishing Company.
- Copeland, D. E., Radvansky, G. A. (2007). Aging and integrating spatial mental models. *Psychology and Aging*, 22(3), 569.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Ferrari, A., & Hagoort, P. (2025). Beat gestures and prosodic prominence interactively influence language comprehension. *Cognition*, 256, 106049.
- Ferreira, V. S. (2019). A mechanistic framework for explaining audience design in language production. *Annual Review of Psychology*, 70, 29-51.
- Galati, A., Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee. *Journal of Memory and Language*, 62(1), 35-51.
- Gerwing, J., Bavelas, J.B. (2004). Linguistic influences on gesture's form. *Gesture*, 4, 157-195.
- Gillespie, M., James, A. N., Federmeier, K. D., Watson, D. G. (2014). Verbal working memory predicts co-speech gesture: evidence from individual differences. *Cognition*, 132, 174–180.
- Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., Wagner, S. (2001). Explaining math: gesturing lightens the load. *Psychological Science*, 12, 516–522.
- Goldin-Meadow, S., Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Reviews of Psychology*, 64, 257–283.
- Göksun, T., Goldin-Meadow, S., Newcombe, N., Shipley, T. (2013). Individual differences in mental rotation: what does gesture tell us. *Cognitive Processes*, 14, 153–162.
- Göksun, T., Özer, D., & Akbıyık, S. (2022). Gesture in the aging brain. In A. Morgenstern, & S. Goldin-Meadow (Eds.), *Gesture in language: Development across the lifespan* (pp. 269–293). De Gruyter Mouton: American Psychological Association.
- Hoetjes, M., Koolen, R., Goudbeek, M., Krahmer, E., Swerts, M. (2015). Reduction in gesture during the production of repeated references. *Journal of Memory and Language*, 79, 1-17.
- Holler, J., Stevens, R. (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology*, 26, 4–27.
- Holler, J., Wilkin, K. (2009). Communicating common ground: how mutually shared knowledge influences the representation of semantic information in speech and gesture in a narrative task. *Language and Cognitive Processes*, 24, 267-289.
- Holler, J., Levinson, S. C. (2019). Multimodal language processing in human communication. *Trends in Cognitive Sciences*, 23(8), 639-652.
- Horton, W. S., Spieler, D. H. (2007). Age-related differences in communication and audience design. *Psychology and Aging*, 22(2), 281.
- Hostetter, A. B., Alibali, M. W. (2007). Raise your hand if you're spatial: relations between verbal and spatial skills and gesture production. *Gesture*, 7, 73–95.
- Iverson, J. M., Capirci, O., Longobardi, E., Caselli, M. C. (1999). Gesturing in mother-child interactions. *Cognitive Development*, 14(1), 57-75.
- Jacobs, N., Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language*, 56(2), 291-303.
- Kandemir, S., Özer, D., Aktan-Erciyes, A. (2023). Multimodal language in child-directed versus adult-directed speech. *Quarterly Journal of Experimental Psychology*, 17470218231188832.
- Kita, S., Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48, 16–32.
- Kita, S., Alibali, M. W., Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Reviews*, 124, 245–266.
- Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of memory and language*, 57(3), 396-414.
- Marstaller, L., Burianová, H. (2013). Individual differences in the gesture effect on working memory. *Psychonomic Bulletin and Review*, 20, 496–500.
- McNeill, D. (2005). *Gesture and thought*. University of Chicago Press.
- Novack, M. A., Goldin-Meadow, S. (2017). Gesture as representational action: a paper about function. *Psychonomic Bulletin and Review*, 24, 652–665.
- Özer, D., Göksun, T. (2020a). Gesture use and processing: A review on individual differences in cognitive resources. *Frontiers in Psychology*, 11, 573555.
- Özer, D., Göksun, T. (2020b). Visual-spatial and verbal abilities differentially affect processing of gestural vs. spoken expressions. *Language, Cognition and Neuroscience*, 35(7), 896-914.
- Özer, D., Özyürek, A., & Göksun, T. (2025). Spatial working memory is critical for gesture processing: Evidence from gestures with varying semantic links to speech. *Psychonomic Bulletin and Review*, 1-15.
- Pouw, W. T., De Nooijer, J. A., Van Gog, T., Zwaan, R. A., Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Frontiers in Psychology*, 5:359.
- Prové, V., Oben, B., Perez, J. (2022). Foreigner directed gesture: larger, faster, longer. *Leuven Working Papers in Linguistics*, 8, 1–31.
- Schubotz, L., Özyürek, A., Holler, J. (2019). Age-related differences in multimodal recipient design: younger, but

- not older adults, adapt speech and co-speech gestures to common ground. *Language, Cognition and Neuroscience*, 34(2), 254-271.
- Shepard, R. N., Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171(3972), 701-703.
- Tellier, M., Stam, G., Ghio, A. (2021). Handling language: How future language teachers adapt their gestures to their interlocutor. *Gesture*, 20(1), 30-62.
- Trafton, J. G., Trickett, S. B., Stitzlein, C. A., Saner, L., Schunn, C. D., Kirschenbaum, S. S. (2006). The relationship between spatial transformations and iconic gestures. *Spatial Cognition and Computations*, 6, 1–29.
- Vigliocco, G., Perniss, P., Vinson, D. (2014). Language as a multimodal phenomenon: implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651), 20130292.