Context-Sensitive Spatial Deixis Through the Lens of Mental Models

Adwitiya Mandal
Department of Computing Science, Umeå University, Sweden

Zoe Falomir
Department of Computing Science, Umeå University, Sweden

Kai-Florian Richter
Department of Computing Science, Umeå University, Sweden

Abstract

The use of spatial deixis, such as *here* and *there*, depends not only on physical distance, but also on how speakers mentally organize the relevant space. This paper examines how contextual factors may influence and shape the use of spatial deixis in an object-selection task. The analysis is based on verbal data from a previous human-robot interaction study, in which 12 participants referred to objects in a shared cupboard environment. Using Norman's distinction of socio-cognitive systems, we interpret the observed deixis behavior as evidence for potential spatial mental models underlying participants' spatial deixis use. We find that speakers' deixis use reflects a combination of absolute, relative, mixed, temporal, and response-biased models rather than a single fixed strategy. Our findings support the view that spatial deixis is context-sensitive, shaped by the mental organization of the space and is variable across individuals, with implications for interpreting spatial deixis in human-robot interaction.

2012 ACM Subject Classification Human-centered computing → Empirical studies in HCI; Human-centered computing → HCI theory, concepts and models; Applied computing → Psychology; Computing methodologies → Cognitive robotics

Keywords and phrases Spatial deixis, here and there, mental models, human-robot interaction, object selection, spatial cognition

Digital Object Identifier 10.4230/LIPIcs.COSIT.2026.24

Category Short Paper

Supplementary Material *Dataset (Excel table)*: https://osf.io/qh5ap/overview

Funding *Adwitiya Mandal*: funded by the Swedish Research Council (VR) under grant 2022-03966.
Zoe Falomir: acknowledges the funding by the Wallenberg AI, Autonomous Systems and Software Program (WASP) awarded by the Knut and Alice Wallenberg Foundation, Sweden.

1 Introduction

Referring to physical space is highly context-dependent [10], ranging from topological relationships between objects to directional spatial relations within larger environments, such as rooms, where other physical entities play a role in defining spatial actions. In all languages, demonstratives play a crucial role in discourse. In English, demonstratives are classified as "proximal" (*this*, *these* and *here*) and "distal" (*that*, *those* and *there*) [1, 12]. They are closely related and are defined by the same deictic roots of joint attention [7, 5], which is considered a fundamental element of social cognition. Among them, spatial deixis (*here* and *there*), are words that depend on the spatial position of the speaker and/or listener and convey location-specific information [16, 9].

The interpretation of spatial deixis depends less on the objective metric space and more on how the speaker conceptualizes the communicative context [8]. Neuro-psychological studies suggest that such spatial distinctions in demonstratives are rooted in embodied interaction

17th International Conference on Spatial Information Theory (COSIT 2026).
Editors: Sabine Timpf, Gabriele Filomena, Armand Kapaj, Rui Zhu, Nicholas Giudice, and Ed Manley;
Article No. 24; pp. 24:1–24:8
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

with one's own space, particularly in the distinction between one's peri-personal and extra-personal space. Peri-personal space is the region of space within a person's reach, where they can interact with objects in their immediate visuo-tactile surroundings by reaching, grasping or touching [2, 3, 13, 14]. In contrast, extra-personal space is the space beyond that where objects are only perceptually visible, but physically inaccessible. The peri-personal space is contextually dynamic. For example, it increases if the person has a tool in their hand. In this context, Skilton suggests that experimental and observational studies highlight these dynamic peri-personal spaces as the "here-space" [19]. These context sensitivity and shared attention perspectives of spatial deixis pose both linguistic and perceptual challenges for robots to understand their human collaborators in an interaction scenario.

Caldano and Coventry [4], building on discussions of the deictic content of demonstratives [19], point out that their usage depends not only on the perceptual space, but also factors, such as object properties, the hearer's position and joint attention. Extending this view, the present paper investigates the research question "how does context influence the use of the spatial deixis 'here' and 'there' (in an object selection task)?"

2 The experimental setting

The present study builds on verbal interaction data collected in a previous Human-Robot Interaction (HRI) experiment [15]. In this experiment, participants were instructed to direct the robot in an imagined scenario of moving objects from a cupboard. It did not actually involve any movement of the robot or the objects. Participants pointed to cups and bowls placed on the cupboard while facing it from the same perspective as the robot, with both remaining at fixed positions through-out the interaction. The study comprised of five different scenarios where the objects were arranged in increasing density and in different configurations, as shown in Fig 1. In each scenario, the robot asked the participant to select an object by both pointing and referring verbally to that object. Participants were instructed to use a verbal expression that at least contained the object class along with a spatial deixis, as they would deem appropriate, for example, "the cup there." These verbal instructions were recorded for all the objects across all scenarios for 12 participants. These data form the basis of the present analysis.

Figure 1 The five experimental scenarios showing the arrangement of objects, adapted from [15].

The recorded data revealed variability in participants' use of spatial deixis, both across individuals and within the same arrangements. This variability suggests that, at least in a multi-object selection task, the flexible nature of the peri-personal space [12] could also be influenced by other cognitive or mental representations to the ones mentioned in the introduction.

Our hypothesis is that within a given environment, the choice between *here* and *there* is shaped not only by metric distance, but also by how users perceive the size, extent and organization of the environment. In the present study, the cupboard serves as a controlled spatial setting for examining such contextual influences (cf. [10]). Although the analysis is grounded in this specific setting, we expect the findings to be transferable to similar spaces and scenarios.

3 Conceptual Framing

To account for the observed variability in deixis use, we examine the data from a socio-cognitive perspective. This framing allows us to consider how the use of *here* and *there* is shaped by the way individuals interpret the interaction situation. Since the boundaries of proximal space are dynamically modulated, participants may draw on embodied spatial perspectives [10], prior experiences, and personal beliefs to form internal mental models for the task at hand and to resolve ambiguities [17, 18]. These mental models are inherently unstable and may be revised by the same individual across different instances of similar tasks.

For our analysis, we adopt Norman's distinction [17] between four related components: the target system $(\mathbf{t})$, the conceptual model of that target system $(\mathbf{C(t)})$, the user's mental model $(\mathbf{M(t)})$ and the researcher's conceptualization of the user's mental model $(\mathbf{C(M(t))})$. Within this framing, the present paper interprets the cupboard space as a structured interaction space, and subsequently uses the participants' observed deixis behavior to construct an analytical account of potential user mental models for the use of spatial deixis. Therefore, in terms of Norman's four components, the experimental setting is as follows:

1. **t:** The target system is the shared cupboard space, which a participant and a robot jointly orient towards from fixed positions with a shared perspective.
2. **C(t):** The conceptual model of the target system is the structured representation of the interaction space. In the study, this consists of the bounded cupboard environment and the spatial arrangements of the objects across the five scenarios, as shown in Fig 1. This allows for analyzing how the use of spatial deixis relates to object placement, relative spatial configurations and contextual variation across scenarios.
3. **M(t):** The participants' mental models are assumed to be shaped by embodied reach, object arrangement, attentional selection [6, 12] and contextual interpretation (see next section).
4. **C(M(t)):** The researchers' conceptualization of the participants' mental models is formulated as a hypothesis about how the cupboard space may be interpreted for deixis use. The cupboard is represented as a 3×3 grid of zones. Since participants approach the cupboard from the left, the left column (zones 1, 4, and 7) is hypothesized to correspond more strongly to *here*, the right column (zones 3, 6, and 9) to *there*, and the middle column (zones 2, 5, and 8) to a potentially ambiguous deixis use. In this way, C(M(t)) provides an initial analytical interpretation of how participants may mentally organize the cupboard into proximal, distal and intermediate spatial regions during deixis use.

The next section presents potential mental models, which may underlie the use of a person's spatial deixis in a given context. These models are treated as components of C(M(t)), without any claims of being exhaustive.

4 Methods and results

4.1 Procedure

Our analysis focuses on the verbal spatial deixis produced by the 12 participants for 174 valid trials (out of 180) in [15]. Participants pointed at an object in each trial while verbally referring to it using spatial deixis (e.g., "the cup *there*)." For each participant, this deixis use was examined at the level of individual objects within each scenario, and subsequently compared across scenarios to trace changes in the use of *here* and *there* across zones and objects. For each object in each scenario, we noted whether it was referred to with *here* or *there*, in which zone of the cupboard it is located (absolute position), its relative closeness to the participant compared to other objects (relative location), and which deixis had been used in references to the object in previous scenarios. We then looked for object-level and scenario-level patterns across participants in order to capture inter-individual variation and to infer possible contextual factors and mental models underlying such variation.

4.2 Analysis

The choice of spatial deixis in a given situation may be shaped by multiple interacting factors, formed consciously or unconsciously. Some of the factors that we consider here are: the **environmental context** (e.g. quantity and arrangement of objects in the cupboard), the **configuration and properties of the objects**, such as relative closeness, and the **human's embodied perspective** relative to the cupboard (e.g. pointing gestures extend participants' peri-personal space). The object size may also influence deixis use, but since the objects in the setup were comparable in size to fit the cupboard, it was excluded from our present analysis.

Considering these factors, we suggest the following potential spatial mental models that enrich **C(M(t))** in the spatial understanding of deixis use:

Absolute model (AM) is characterized as being driven by the contextual environment. In the case of the cupboard setting, the spatial layout itself provides a structured frame that can bias deixis selection as mentioned in C(M(t)) in Section 3.

Relative model (RM) is defined in terms of the relative arrangement of objects within a single scenario. Rather than interpreting an object solely by its absolute proximity to the participant, it examines whether the use of *here* or *there* depends on how that object is positioned relative to other objects in the same arrangement. In this way, RM captures how relational spatial context within a scene may shape deixis use.

Mixed model (MM) is conceptualized as an integration, or interaction, of both absolute and relative mental models.

Temporal model (TM) captures possible carry-over effects in deixis use for an object across scenarios. It examines whether the deixis assigned to the same object in one arrangement is influenced by how that object was interpreted in preceding arrangements. This can be defined as an effect of mental conditioning across scenarios or an adaptability of the deixis use across arrangements.

Response-biased model (RbM) is defined in terms of individual response bias. A participant is considered response-biased if they consistently favor one deixis, for a majority of the objects within a scenario and across scenarios, independent of their absolute and relative positions. Predominant use of *here* might suggest proximal contextual space, within their arm's reach, whereas predominant use of *there* might suggest a more distal space.

4.3 Results

Table 1 shows the applicability of the suggested spatial mental models for each participant (P_id) for all scenarios. Based on these observed patterns, we classify the participants into four groups that characterize their dominant deixis use strategies.

Structured strategy. The first group shows a relatively structured use of spatial deixis across scenarios. These participants primarily seem to follow AM, while also adapting their responses according to the relative positions of nearby objects. This pattern is mostly

Table 1 The table shows the mental models that fit the participants' spatial deixis use in and across scenarios 1 to 5. A tick indicates that the respective mental model may explain deixis use.

P_id.	Mental Models	Scenarios				
		1	2	3	4	5
1	AM	✓	✓	✓	✓	✓
	RM			✓	✓	✓
	MM			✓	✓	✓
	TM		✓	✓	✓	✓
	RbM					
2	AM		✓	✓		
	RM	✓	✓	✓	✓	
	MM		✓	✓		✓
	TM		✓	✓		
	RbM				✓	✓
3	AM	✓	✓	✓	✓	
	RM			✓	✓	
	MM			✓	✓	
	TM		✓	✓	✓	✓
	RbM					✓
4	AM		✓		✓	✓
	RM	✓	✓	✓	✓	✓
	MM		✓		✓	✓
	TM		✓			✓
	RbM					
5	AM	✓	✓			
	RM		✓	✓	✓	
	MM		✓	✓		
	TM		✓		✓	
	RbM				✓	✓
6	AM	✓	✓	✓	✓	
	RM		✓	✓	✓	
	MM		✓	✓	✓	
	TM		✓		✓	
	RbM					
7	AM	✓	✓	✓	✓	✓
	RM		✓	✓	✓	✓
	MM		✓	✓	✓	✓
	TM		✓		✓	✓
	RbM					
8	AM	✓	✓	✓	✓	
	RM		✓	✓	✓	
	MM		✓	✓	✓	
	TM				✓	
	RbM					✓
9	AM	✓	✓	✓	✓	✓
	RM		✓		✓	
	MM		✓		✓	
	TM		✓	✓	✓	✓
	RbM					✓
10	AM	✓	✓	✓		
	RM			✓		
	MM			✓	✓	✓
	TM		✓	✓	✓	✓
	RbM		✓		✓	✓
11	AM	✓	✓	✓	✓	✓
	RM		✓	✓	✓	✓
	MM		✓	✓	✓	✓
	TM				✓	
	RbM					✓
12	AM	✓			✓	
	RM		✓	✓	✓	✓
	MM		✓	✓	✓	
	TM		✓			✓
	RbM			✓		✓

observed for participants 1, 3, 7, 9, and 11. For both participants 1 and 3, the deixis assigned to cup_2 (the cup in zone 2 of the cupboard) remained stable across scenarios, suggesting also a TM effect.

Relative strategy. The second group relies more strongly on RM, where the deixis assigned to an object depends on its relative position to neighboring objects rather than its absolute position on the cupboard. This pattern is clearest for participant 4 but is also observed for participants 7, 8, 11 and 12.

Hybrid strategy. The third group reflects the use of MM, combining both absolute and relative interpretations. This pattern is particularly visible for participants 7 and 11. The responses of participants 2 and 4 suggest sensitivity both to the relative arrangement of objects and to the height of the cupboard. For participant 2, this pattern becomes less stable in scenario 5, whereas participant 4 shows a more structured use even with increased object density.

Variable strategy. The final group shows more inconsistent patterns in their use of spatial deixis across scenarios. For participant 5, only limited local structure seems to be present in scenarios 1 and 2, while participant 10 and 12 shows no clear regularity at all across the scenarios.

In addition, we find the use of RbM in the more complex scenarios 4 and 5. In scenario 5, with increased object density, participants 2, 3, 5, 8, 9, 10, 11, and 12 show a more biased response towards either *here* or *there*, with their use less linked to absolute or relative position of the objects.

5 Discussion

The results suggest that the use of spatial deixis in the cupboard setting cannot be explained by a single stable principle alone. Although the initial hypothesis in C(M(t)) proposes a broad proximal-distal organization of the cupboard as has been found in previous research [4], the analysis of the verbal data indicates that participants do not just rely on this spatial partition in a multi-object arrangement. Instead, their deixis use appears to reflect multiple interacting influences, including absolute position, the relative arrangement of the objects and, in some cases, biases over a specific deixis. The present results differ from previous results in the literature (e.g., [4]), achieved in single-object settings. They support the Norman-inspired framing [17] adopted in this paper, where the participants' latent mental model M(t) may be distinctively different from the experimenters' conceptual model C(M(t)).

From a linguistic perspective, the variability is also consistent with the fact that *here* is not a uniformly simple proximal term. As noted by Fretheim et al. [11], *here* may function as a token-reflexive, anaphoric or deictic expression and its interpretation depends on the contextual information available to the addressee. In this sense, the present analysis is consistent with the view that the use of *here* and *there* is shaped not only by physical metric distances, but also by how the scene is mentally and interactionally organized. The participant groupings further support this interpretation, where participants employed different strategies for their deixis use. Scenario 1 shows the clearest pattern, with 10 out of 12 participants following AM. This suggests that in simpler spatial configurations, people tend to rely on the overall structure of the environment. In contrast, Scenarios 2 to 5 show a combination of different mental models, indicating that increased object density leads to more flexible, context-dependent and less stable deixis use.

The present study offers an initial socio-cognitive interpretation of the observed variability of spatial deixis. Still, there is a need for a more systematic exploration of potential context factors influencing spatial deixis use and data collection from more participants and in other, more systematically selected environment settings in future research. Still, it appears that together, these findings support the idea that the participants' (or speakers') mental model M(t) is flexible and context-sensitive and that the experimenters' (or recipients') mental model C(M(t)) must therefore accommodate more than one strategy for the interpretation of spatial deixis.

This is particularly relevant for designing "artificial" recipients, such as social robots, and for human-robot interaction, where predefined spatial zones may fail to capture the variability in human spatial language, especially when object arrangements are dense or ambiguous. Overall, our study supports a view of spatial deixis being shaped as a combination of contextual and individual factors, mentally organized and variable across individuals.

References

1. Hissah Nasser Alothman and Haroon N Alsager. Acquisition of l2 english spatial deixes by arabic-speaking children. *Frontiers in Psychology*, 13:997110, 2022.
2. Rory J. Bufacchi and Gian Domenico Iannetti. An action field theory of peripersonal space. *Trends in Cognitive Sciences*, 22(12):1076–1090, 2018. `doi:10.1016/j.tics.2018.09.004`.
3. Karl Bühler. *Sprachtheorie*, volume 2. Jena Fischer, 1934.
4. Michela Caldano and Kenny R. Coventry. Spatial demonstratives and perceptual space: To reach or not to reach? *Cognition*, 191:103989, 2019. `doi:10.1016/j.cognition.2019.06.001`.
5. Malinda Carpenter and Michael Tomasello. Joint attention and imitative learning in children, chimpanzees, and enculturated chimpanzees. *Social Development*, 4(3):217–237, 1995.
6. Marvin M Chun, Julie D Golomb, and Nicholas B Turk-Browne. A taxonomy of external and internal attention. *Annual review of psychology*, 62(1):73–101, 2011.
7. Holger Diessel and Kenny R Coventry. Demonstratives in spatial language and social interaction: An interdisciplinary review. *Frontiers in Psychology*, 11:555265, 2020.
8. Nick J Enfield. Demonstratives in space and interaction: Data from lao speakers and implications for semantic analysis. *Language*, 79(1):82–117, 2003.
9. Charles J Fillmore. Santa cruz lectures on deixis, 1971. *(No Title)*, 1975.
10. Christian Freksa, Alexander Klippel, and Stephan Winter. A Cognitive Perspective on Spatial Context. In Anthony G. Cohn, Christian Freksa, and Bernhard Nebel, editors, *Spatial Cognition: Specialization and Integration*, volume 5491 of *Dagstuhl Seminar Proceedings (DagSemProc)*, pages 1–16, Dagstuhl, Germany, 2007. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/DagSemProc.05491.4`.
11. Thorstein Fretheim, Nana Aba Appiah Amfo, and Ildikó Vaskó. Token-reflexive, anaphoric and deictic functions of "here". *Nordic Journal of Linguistics*, 34(3):239–294, 2011. `doi:10.1017/S0332586512000030`.
12. Harmen B Gudde, Jacqueline Collier, and Kenny R Coventry. Spatial demonstratives and perspective taking in english and japanese. *Cognitive Science*, 50(3):e70183, 2026.
13. Edward T Hall. Proxemics—the study of man's spatial relations. *Man's image in medicine and anthropology*, 4:1003–1026, 1963.
14. William F Hanks. *Referential practice: Language and lived space among the Maya*. University of Chicago Press, 1990. `doi:10.1017/S0047404500017036`.
15. Adwitiya Mandal, Kai-Florian Richter, and Zoe Falomir. How robots understand 'here' and 'there': A perceptual model for spatial deixis. In *Companion Proceedings of the 21st ACM/IEEE International Conference on Human-Robot Interaction*, HRI Companion '26, pages 1018–1022, New York, NY, USA, 2026. Association for Computing Machinery. `doi:10.1145/3776734.3794548`.

16. Sandra Murphy. Children's comprehension of deictic categories in oral and written language. *Reading Research Quarterly*, pages 118–131, 1986.
17. Donald A Norman. Some observations on mental models. In *Mental models*, pages 7–14. Psychology Press, 2014.
18. Jane Oakhill and Alan Garnham, editors. *Mental models in cognitive science: essays in honour of Phil Johnson-Laird*. Psychology P., 1996.
19. Amalia Skilton. The deictic content of demonstratives. *Language and Linguistics Compass*, 18(4):e12519, 2024. doi:10.1111/lnc3.12519.