Experimental perspectives on spatial deictic expression acquisition in Thai[1]

Nattanun CHANCHAOCHAI — *Department of Linguistics and Center of Excellence in Southeast Asian Linguistics, Faculty of Arts, Chulalongkorn University*
Florian SCHWARZ — *Department of Linguistics, University of Pennsylvania*

Abstract. This study explored the role of perspectives on the interpretation of Thai spatial deictic terms *lǎŋ* 'behind' and *nâː* 'in front of.' We designed an experiment where the relative location of the experimenter and the participant differed across two sessions. Utilizing 5 objects of 5 different colors arranged in a row, the experimenter asked the participant '*The pen/balloon of which color is behind/in front of the x pen/balloon?*,' where *x* is the color of one of the middle three items. Participants were native speakers of Thai, including 31 adults, 60 typically developing children, and 30 children with Autism Spectrum Disorders (ASD). We found that for Thai-speaking adults, the convention for *behind/in front of* seems to involve object-based allocentric viewpoint, relativizing the deictic notions as if the objects were facing themselves. However, the relative location of the experimenter hugely affected the adults who started off sitting opposite to the experimenter. In contrast, children with TD, egocentric interpretations, projecting their own front and back onto the objects, are preferred for non-fronted, visible objects. Verbal questions and mere presence of Speaker, i.e., without gazes nor gestures, are enough to introduce ambiguities between FoRs for adults and children with TD. Despite their differences from adults in overall perspective resolution, TDs do take into account the experimenter's viewpoint as a contextual resolution option. On the other hand, children with ASD were not affected by experimenter's perspective, and lacked a clear contrast between the opposite deictic terms. They also seem to have general difficulties in grasping the basic relational nature of the two deictic terms.

Keywords: spatial deixis; perspective-taking; relative frame of reference, contextual resolution; autism spectrum disorders.

1. Introduction

Spatial deixis involves resolution of a complex context dependence. While demonstratives like *this* and *that* are mostly anchored to the speaker, this is not always the case for terms such as *behind* and *in front of*. To decide what is *behind* and *in front of* something depends on several factors, including 'frontedness' of entity, additional perspectives or viewpoints of another participants, non-linguistic cues, e.g., gazes, gestures, and developed conventions of different languages. For entities that have obvious fronts, such as dolls or cars, the expressions are usually anchored to the entities. For non-fronted entities, such as balls or pens, the other factors are more likely to come into play. The speaker may choose their own viewpoint or another participant's viewpoint, if there is any, as well as taking into consideration the developed convention of their language. The hearer of such terms may additionally take on non-linguistic cues given by the speaker. We report an experiment investigating factors at play in making these choices in Thai. Given the crucial role of consideration of different perspectives, we look across populations generally assumed to differ in their resources for this, namely children with

[1]The travel to the conference was supported by Chulalongkorn University. This project would have been impossible without the help of all the teachers, parents, and children at Kasetsart University Laboratory School. We would also like to thank everyone in the Experimental Study of Meaning Lab, the Embick Lab, and audiences at SuB27 and XPRAG 2022 for valuable comments and discussion.

typical development (TD), children with autism spectrum disorders (ASD), and adults. Prior work found the terms *behind/in front of* to respectively refer to objects *hidden/visible* relative to their body in TDs (Johnston 1984). In ASDs, deictic terms have been found to pose additional challenges with regards to both person deixis (Bartak and Rutter 1974; Charney 1980: a.o.) and spatial deictic terms and gestures (Loveland and Landry 1986; Hobson et al. 2010: a.o.).

The present paper begins with the background literature on spatial deixis, its acquisition, and deixis and autism (Section 2). Section 3 presents the methods of the study. Sections 4 and 5 describe and discuss the results of the experiment.

2. Background

2.1. Spatial deixis and frames of reference

Deixis serves as a linguistic hook into the contextual, perspectival aspects of utterances. It refers to kinds of references that require contextualization based on certain discourse elements, including person deixis (e.g., *I* and *you*), time/temporal deixis (e.g., *now* and *then*), place/spatial deixis (e.g., *here* and *there*), discourse deixis (e.g., *this* and *that*), and social deixis (e.g., honorifics such as *du* vs *sie* in German) (Fillmore 1971, 1975; Lyons 1977; Levinson 1983). Deictic information is important for interpreting utterances. When such information is lacking, a totally unanchored message cannot be fully interpreted, as seen in example (1) of a message that is accidentally come across via a bottle afloat in the sea.

(1) Meet me here at noon tomorrow with a stick about this big. (Fillmore 1971: p. 39)

To anchor a deictic expression, a *deictic center*, also known as 'origo' (Bühler 2011; Diessel 2014), is needed. By default, the deictic center is assumed to be the speaker, but in certain contexts, it can be 'shifted'. For instance, while the spatial deictic term such as *come* usually describes motion towards the deictic center (Talmy 1975; Oshima 2006; Wilkins and Hill 1995), i.e., the speaker in default cases, deictic center may be shifted to the hearer (2a) or another entity altogether (2b).

(2) a. Can I come visit you?
b. John was preparing a meal. Then, the cat came to him. (Oshima 2006: p. 287)

Such deictic projection, where the deictic center/origo is projected from the speaker to the hearer or others, is called 'deictic-center shifting' (DCS; Levinson 1983; Fillmore 1997). It has been found to be important for the analysis of demonstrative reference. In contexts where there is a direct collaboration between interlocutors, proximal or distal linguistic expressions were found to be interpreted based on the partners' 'action space,' as opposed to the speakers' own action space (Rocca et al. 2019). Similarly, for other spatial terms, such as *to the left/right; in front of*, the presence of another person gazing or reaching for tested objects induced more deictic projections where respondents took the other person's perspective, instead of their own (Tversky and Hard 2009; Tosi et al. 2020).

Spatial deictic expressions vary across languages. Different languages make use of different frames of reference (FoR), i.e., the underlying coordinate system for locating a reference object. One important FoR distinction in the developmental and behavioral psychology and neuroscience literature (e.g., Paillard (1991); Burgess et al. (1999); see Levinson (2003: pp. 28-29) for discussion) is the distinction between 'egocentric' vs. 'allocentric' (non-egocentric). With an egocentric FoR, the meanings of expressions are anchored to one's own perspective, e.g., 'behind' for one own's back, whereas they are anchored elsewhere with an allocentric FoR, e.g., others' or object's back. In addition to such distinction, based on the cross-linguistic data from the Max Planck Institute for Psycholinguistics at Nijmegen, Levinson (2003) proposed a tripartite taxonomy of FoR for world's languages, mainly characterizing the relations between three entities, including a 'figure' entity, i.e., the entity to be located, a 'ground' entity, i.e., the entity that is referenced, and a 'coordinate', i.e., the entity that is the deictic center/origo of the coordinate system. Languages vary in their preference or availability of the three types of FoR.

An absolute frame is an environment-based, i.e., geocentric, frame with infinite region of space covered by the frame. For instance, in Figure 1a, the figure (crab) is related to the ground (chicken) via the deictic center (hill). The entire region that is uphill from the chicken is available to be covered by the frame. An intrinsic frame, on the other hand, only involves a figure and a ground, as the latter also serves as the deictic center (origo). In Figure 1b, the frame only relates the figure (crab) to the ground (Sue), which is also the deictic center. The picture is more complicated with a relative frame, where the ground is not the same entity as the deictic center. As seen in Figure 1c, an entity's viewpoint needs to be chosen as the origo of the sentence. If an egocentric FoR is at play, the speaker would be the deictic center. The term 'behind' would then be interpreted as towards the same direction as the speaker's back, making the figure be the orange pen. Conversely, if the speaker makes use of an allocentric FoR, with the hearer being the origo, it is the blue pen which would be 'behind' the purple pen, i.e., locating towards the same direction as the hearer's back with reference to the purple pen.

While intrinsic FoRs are typologically available to all languages, relative FoRs are not. Shusterman and Li (2016) proposed possible explanation for this. First, relative FoRs are harder than intrinsic FoRs conceptually, because of how the deictic center may need to be shifted from the ground. Secondly, relative FoRs allow for ambiguity to arise, further contributing to their difficulty, as seen in Figure 1c. To account for such ambiguity, different languages adopt different developed conventions as the default interpretation of spatial deictic expressions. However, the convention within the same language may also vary by term. For instance, an egocentric viewpoint might be used for the terms *left/right* but an allocentric viewpoint, sometimes of an imaginary listener's, is used for *front/back* (Levinson 2003; Shusterman and Li 2016). Hausa, Tamil (at least the NaTar caste, Ramnad district dialect), and English were used to instantiate these different conventions in (Levinson 2003: pp. 87-88) and (Shusterman and Li 2016: pp. 119-120). When a ground object is non-fronted, Hausa speakers project their egocentric coordinates onto the object for both front/back and left/right. Tamil's convention, on the other hand, is the total opposite, making use of the allocentric viewpoint for all the terms. This is different from the English, where the coordinate system is mixed, i.e., an egocentric projection for left/right but an allocentric frame for front/back. Since Thai is a language with a relative frame of reference, ambiguity arises, and perspectives play an important role in the interpretation of spatial deictic expressions. Which pattern is preferred by Thai-speaking adults, children with TD, and children with ASDs is the interest of our paper.

2.2. The acquisition of spatial deictic terms

Clark (1978) claimed that deictic gesture provides a basis for children's acquisition of verbal

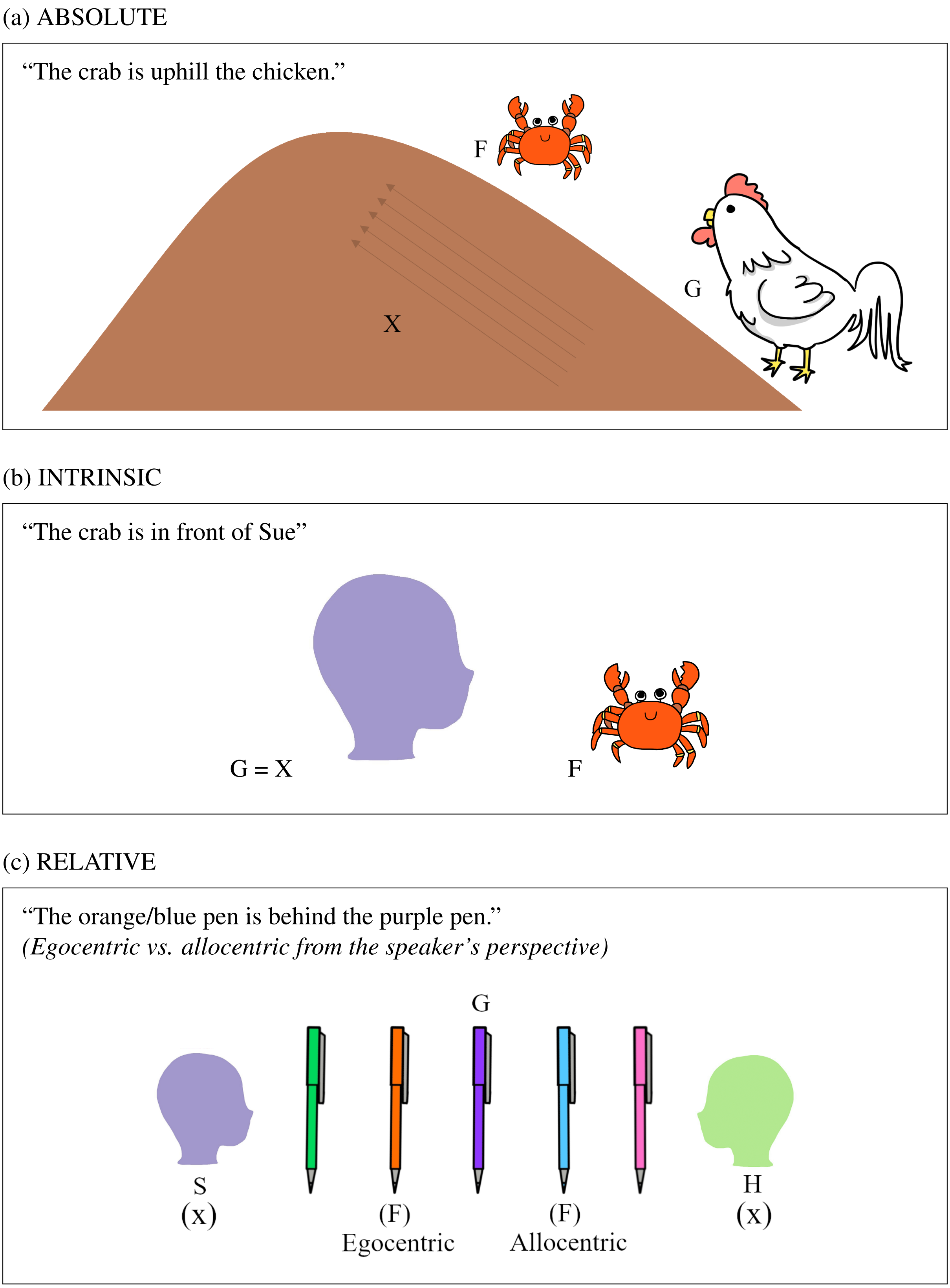

Figure 1: Absolute (a), intrinsic (b), and relative (c) tripartite typology. (F: figure; G: ground; X: coordinate, deitic center/origo)

deixis and that demonstratives are used very frequently in children's early years, being among children's first 50 words. Although there is little empirical research on the acquisition of spatial deixis, recent studies have found that the claim needs to be revisited and verified. Diessel and Coventry (2020) looked at the data in the CHILDES corpora from English (N = 10), Dutch (N = 3), Hebrew (N = 4), and Japanese (N = 3) and found large mean proportions of demonstratives in early child speech ranged from 5.88% to 8.27%. However, the strong claim that demonstratives are within children's first 50 words were not verified. González-Peña et al. (2020) used the language production data of 18- to 24-month-old Spanish- and English-speaking children from the CHILDES corpora (N = 66) and McArthur-Bates Communicative Development Inventories database (N = 950) to explore the long-standing claim and found that demonstratives are not typically among the first 50 words of children.

In learning a relative language, Piaget (1928: pp. 107-108) argued that children initially map FoR terms such as left and right to their own body (Stage I: age 5-8) before considering the interlocutor's point of view (Stage II: age 8-11) and eventually being able to consider FoR from the object's point of view (Stage III: age 11-12). These stages correspond to Piaget's (1928) social stages where ego-centrism begins to decline around the age of 7-8. While children may start to use FoR terms early, the developmental time course before these terms can involve other people's viewpoints is protracted (Shusterman and Li 2016). Additionally, cultural and language-specific influences may come into play, and different spatial terms may also have different acquisition patterns.

While English-speaking children struggle to take an allocentric viewpoint for the terms *left-right*, failing to talk about the *left* and *right* sides of dolls, the non-egocentric interpretation of *back/front* was found to be available to TDs for fronted objects such as dolls (Shusterman and Li 2016). The fact that these two terms, *back/front*, were also found to respectively refer to objects *hidden/visible* in children with TD (Johnston 1984) may ease the children's adoption of a novel viewpoint that may not correspond to their own. At the age of 4, Shusterman and Li (2016) found that English-speaking children with TD equally chose both egocentric and geocentric interpretations for *back/front*.

With regards to children with ASD, social deficits, language and communication deficits, and repetitive behaviors are the three core clinical features of ASD, with pragmatics and discourse deficits being generally accepted to be central to language deficits in autism (for reviews, see (Lord and Paul 1997; Tager-Flusberg 1999; Wilkinson 1998). Among pragmatic deficits, children with autism are known to have difficulties with person deixis (see, for instance, Bartak and Rutter 1974; Charney 1980; Chiat 1982; Fay 1979; Loveland 1984). Spatial deitic terms and gestures were also found to pose challenges to children with autism (Loveland and Landry 1986; Hobson et al. 2010).

With all the factors, including frontedness of entity, additional perspectives or viewpoints of another participants, non-linguistic cues, e.g., gazes, gestures, and developed conventions of different languages, it is of no surprise that the acquisition time course for spatial deictic terms such as *behind* and *in front of* would be protracted for children with TD and children with ASD. However, while, as mentioned earlier, *behind* and *in front of* seem to allow for a non-egocentric interpretation for fronted objects earlier in the acquisition time course than *left* and *right*, it is still unclear whether and how an additional perspective of another participant would affect children's interpretation. On top of that, if other factors are controlled for, i.e., tested with non-fronted objects and without non-linguistic cues, would they only interpret the terms based on the developed convention of their language (or their own generalization) or also allow the other person's perspective as a viable option for their interpretation as well?

3. Methods

Our experiment tests the interpretations of Thai spatial deictic terms *lǎŋ* 'behind' and *nâ:* 'in front of', while varying the relative location of experimenter (E) and participant (P) (see Figure 2) across two sessions (administered on two different days or in the morning and in the afternoon). Scenario 1 had E and P sitting on the same side of the table, while they sat on opposite sides in Scenario 2. The shared perspective in Scenario 1, where relativization to E and P yield equivalent interpretations, simplifies contextual resolution choices, whereas Scenario 2 is more challenging. Order of scenarios was counterbalanced across participants, with Group A in Scenario 1 first, and Group B Scenario 2. The setup utilized 5 different color pens and balloons respectively.

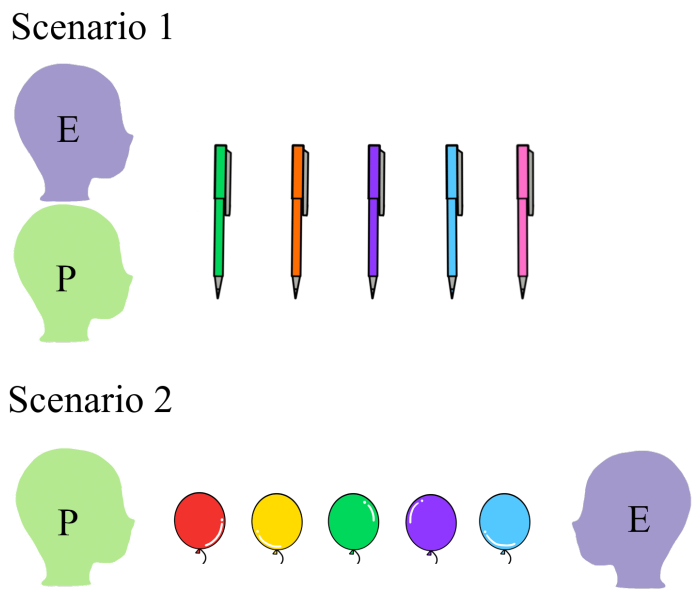

Figure 2: The two scenarios in the experiment (E: Experimenter; P: Participant).

Before the main experiment participants had to name the color of each item to ensure basic understanding. The experimenter then arranged the items in a row and asked the question below (3), with *y* being the color of one of the middle three items. There were 12 trials, with items being rearranged every 3 trials. The experimenter only looked at the data collection form to avoid giving any non-linguistic cues of gazes and movements.

(3) pà:k-ka:/lû:k-pò:ŋ sǐ: ʔa.raj jù: k^hâ:ŋ-nâ:/k^hâ:ŋ-lǎŋ pà:k-ka:/lû:k-pò:ŋ sǐ: **y**
pen/balloon color what COP in front of/behind pen/balloon color **y**
'The pen/balloon of which color is in front of/behind the pen/balloon of **y** color?'

3.1. Participants

Native Thai speaking participants included 31 adults (*M* age = 37.52) and two groups of children. Children with ASD (*N* = 30; *M* age = 9;6) and their typically-developing controls (TDs; *N* = 60; *M* age = 7;11) were recruited from Kasetsart University Laboratory School, Center for Educational Research and Development. All participants from both years had normal hearing and normal or corrected-to-normal vision. The Ravens Standardized Progressive Matrices (Raven et al. 2000) were administered to both groups of children for the assessment of nonverbal IQ (NVIQ; ASD *M* = 94.74; TD *M* = 116.35). The scores were converted using the norms in the 1979 British Standardisation of the Standard Progressive Matrices (Raven 2000, pp. 39-40). Children in both groups had normal hearing and normal or corrected-to-normal vision. This study was approved by the Institutional Review Board at the University of Pennsylvania. Having been informed about the study and their rights, the parents of all the participants provided written consent for their child to participate in the study.

	Adults			ASD			TD		
Group	A	B	Total	A	B	Total	A	B	Total
N (Female *N*)	15(9)	16(11)	31(20)	14(2)	16(1)	30(3)	28(4)	29(6)	57(10)
$M(SD)$ Age	34.67 (6.54)	40.19 (8.55)	37.52 (8.02)	9;6 (1.98)	9;6 (2.13)	9;6 (2.03)	7;10 (1.98)	8;0 (1.86)	7;11 (1.91)
$M(SD)$ NVIQ	NA	NA	NA	105.03 (27.27)	85.74 (17.18)	94.74 (24.12)	118.31 (18.47)	114.47 (18.26)	116.35 (18.30)

Table 1: Participant information

3.2. Data analysis

Responses were coded in terms of distance from the participant, not taking into account where the experimenter was, with Position 0 being the position of the ground (object of **y** color), Position 1 being the position of the pen/balloon adjacent to the ground object that is one position further from the participant, and Position -1 being the position of the pen/balloon adjacent to the ground object that is one position closer to the participant. Figure 3 illustrates (1) how responses were coded for their position indices (Position -2, -1, 0, 1, or 2) and (2) for the term *behind*, which position is considered egocentric vs allocentric interpretation from the participant's perspective in this paper. This is in contrast to the FoR term *in front of* where a egocentric interpretation applies to Position 1, and a allocentric interpretation to Position -1.

Mixed Effects Logistic Regression Models were run using the lme4 package (Bates et al. 2015: version 1.1-28) in the R environment (R Core Team 2016: version 4.1.2) to test whether participants' responses follow the hearer-allocentric pattern of choosing Position 1 for *behind* and Position -1 for *in front of* or not. Main effects included `TrialNumber` and the interactions between `Scenario` (Scenario 1 or 2), `ScenarioOrder` (Scenario 1 first or not), and `ParticipantGroup` (adults, children with TD, children with ASD), and a random effect for participants. `ParticipantGroup` was treatment coded to set different reference levels for each participant group.

To model errors in both groups of children, another mixed effects logistic regression model was run to test whether the children gave strictly incorrect responses (i.e., *z* or objects non-adjacent

"Behind the purple pen."
(Egocentric vs. allocentric from the hearer's perspective)

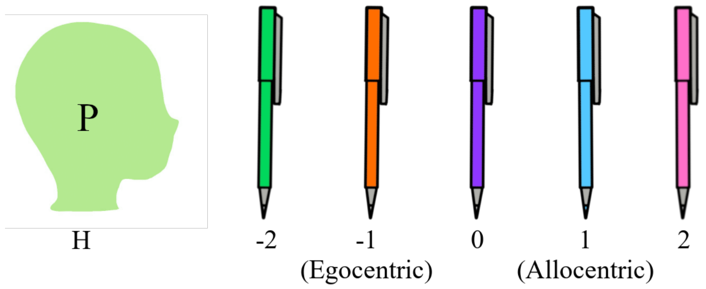

Figure 3: Example coding for responses to the question to locate a pen *behind* the ground (purple pen). (P: Participant; H: Hearer).

to z) or not. Main effects included `Scenario`, `ScenarioOrder`, `Age`, `NVIQ`, `TrialNumber`, the interactions between `Terms` (*in front of* or *behind*) and `ParticipantGroup` (children with TD or children with ASD), and a random effect for participants. `ParticipantGroup` was treatment coded to set different reference levels for each participant group.

4. Results

Overall pattern Figure 4 shows the overall results across participant groups collapsing over scenarios and groups (scenario order). Overall, the adults tended towards choosing Position 1 for *behind* and Position -1 for *in front of*, preferring an allocentric interpretation (e.g., *behind* z = z's back). Children with TD show the opposite pattern, suggesting an egocentric interpretation, relative to their own perspective (e.g., *behind* x = towards the same direction as P's back). Children with ASD generally tended towards objects closer to them (Position -1) regardless of deictic term (*in front of* or *behind*).

Perspectives and Contextual Resolution Both the Scenario ($\beta = 1.54, p < 0.01$) and Scenario order ($\beta = 3.74, p < 0.01$) factors drastically modulated response patterns for adults, as shown in Figure 7. The adults in Group A, who first shared perspective with E, strongly followed the allocative pattern above, although to lesser extent in their second session with divergent perspectives for P and E. In contrast, adults in Group B, who start across from E in Scenario 2, exhibit no overall interpretation preference, suggesting an even mix of contextual resolution choices across speakers. When subsequently sharing perspective with E, their pattern is the exact opposite of Group A, suggesting prominent use of interlocutors' perspective.

For children with TD, the response pattern did not vary across the participants in Group A and B ($\beta = 1.6, p = 0.13$). However, Scenario did affect their response pattern. The TD children were found to perform significantly differently in Scenario 1 versus 2 ($\beta = 0.96, p < 0.01$), with Scenario 1 (sitting together) yielding significantly more allocentric responses. This suggests that children with TD consider E's perspective as a possible viewpoint for the resolution.

While children with TD exhibited similar shifts in performance across sessions, especially

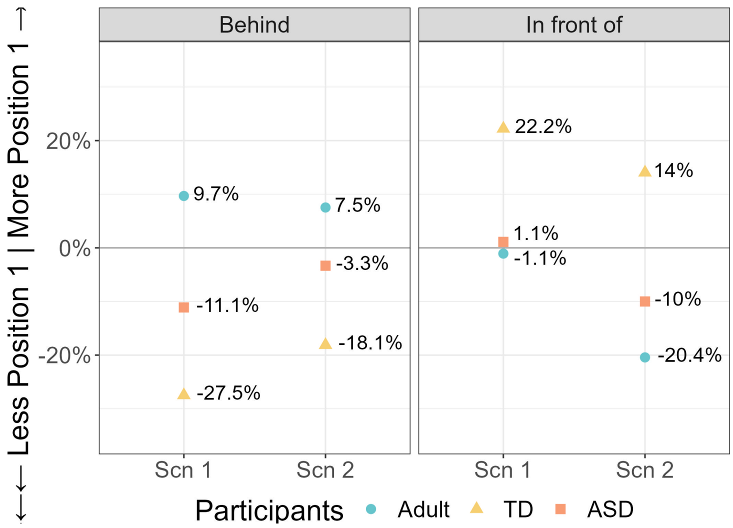

Figure 4: The percentage differences between choosing the object in Position 1 and Position −1. Positive differences indicate the overall preference of the Position 1 choice.

for Group B, no scenario ($\beta = -0.54, p = 0.25$) nor scenario order ($\beta = -0.11, p = 0.87$) effects were found for the ASD group, as seen in Figure 8 and 9. In addition to their minimal performance shifts across sessions, children with ASD gave strictly incorrect responses (e.g., x or objects non-adjacent to x) significantly more often than children with TD (behind: $\beta = 4.62, p = 0.005$; in front of: $\beta = 3.70, p = 0.02$). This can be seen from Figure 5, which presents the overall results across groups, conditions, and sessions and Figure 6, which collapsed the overall rates for choosing the ground object (at Position 0) are presented for each group of participants.

The shifts in performances of Group B between sessions were not significantly different between adults and children with TD ($\beta = 1.03, p = 0.26$) but were significantly different between adults and children with ASD ($\beta = 2.71, p < 0.01$) and marginally significant between children with TD and children with ASD ($\beta = 1.69, p = 0.06$).

5. Discussion

When the initial scenario makes the contrast between speaker's and hearer's perspective moot, adults in Group A employ an object-based allocative viewpoint, with the object construed as facing them. But when distinct perspectives are initially in play (and an allocative viewpoint could be construed as object facing speaker or hearer), responses by adults in Group B suggest an even distribution across contextual resolution options. Once E and P's perspective align in Session 2, interlocutor-based interpretations (E or P, equivalently) dominate the response pattern, in contrast to Group A's allocative viewpoint. Such interpretations relativized to the participant also dominate the response patterns of the children with TD overall, in contrast to Johnston (1984)'s findings ('visible' = *in-front-of* vs. 'hidden'*behind*, which corresponds to allocative resolution). However, in our scenarios, objects were not actually hidden from sight, which plausibly makes an allocative viewpoint less salient, leading the children with TD to prefer an egocentric viewpoint, in line with tendencies based on availability of theory of mind reasoning at their age.

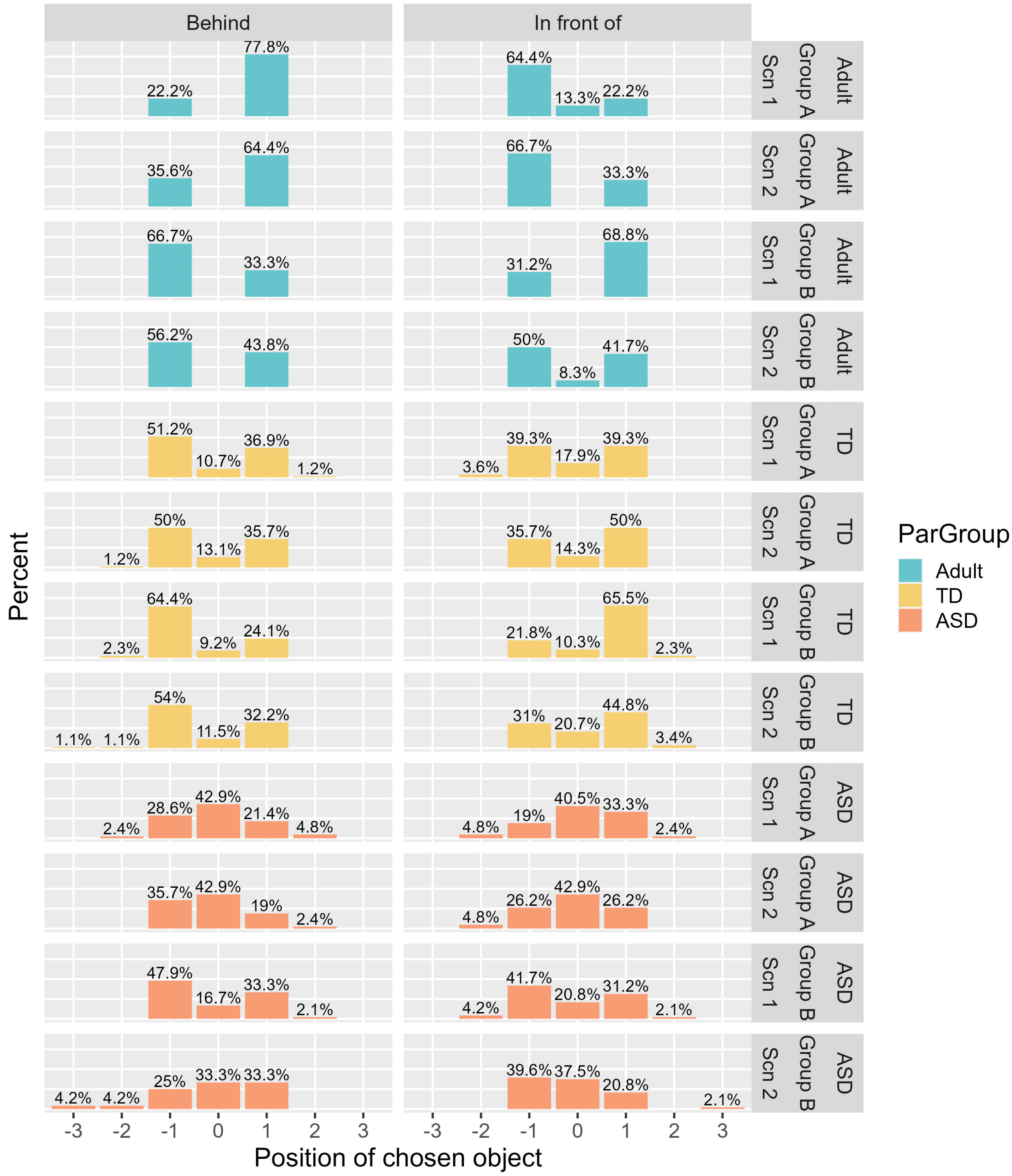

Figure 5: The results in *behind* and *in front of* conditions across participant groups, groups, conditions, and sessions.

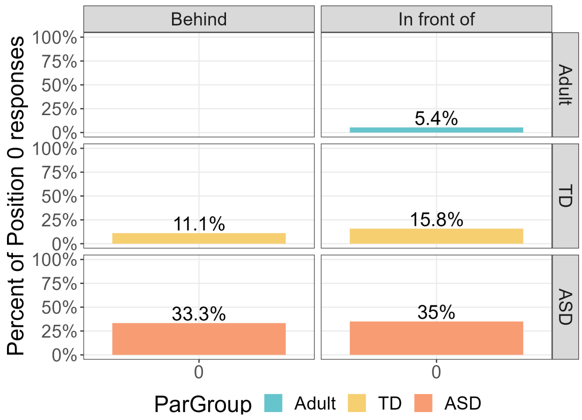

Figure 6: The percentages of choosing the object in Position 0 by term and by participant group.

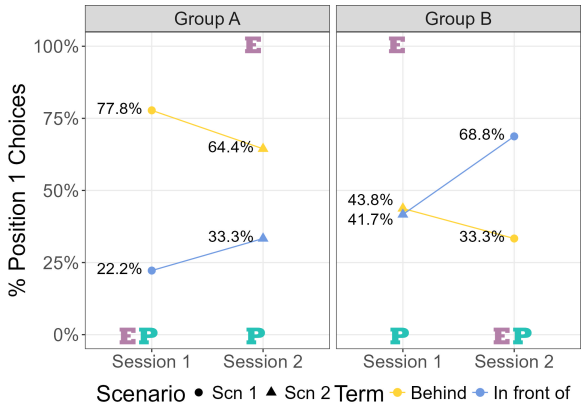

Figure 7: The percentages of responses by the adults in Group A and Group B by preposition, by session, and by location of experimenter.

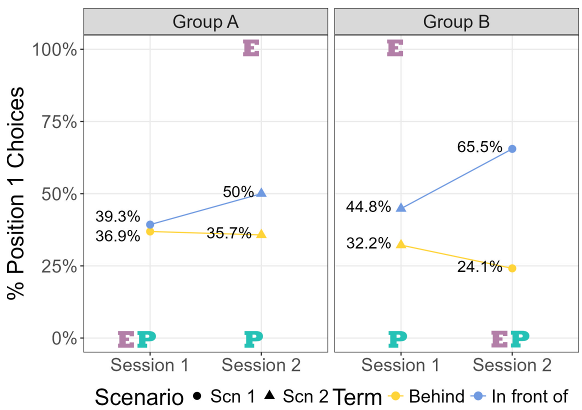

Figure 8: The percentages of responses by the children with TD in Group A and Group B by preposition, by session, and by location of experimenter.

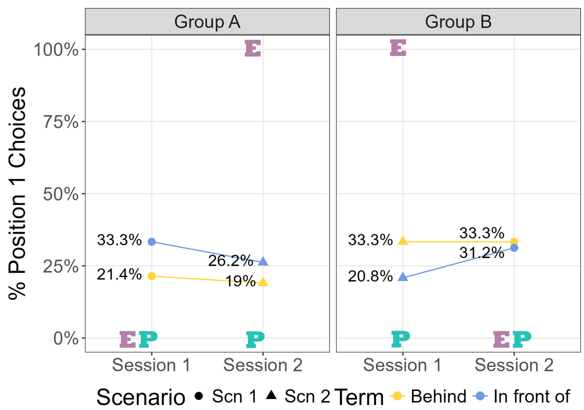

Figure 9: The percentages of responses by the children with ASD in Group A and Group B by preposition, by session, and by location of experimenter.

Both children with TD and adults were similarly affected by the change in experimenter's perspective in Group B, suggesting that despite their differences in overall perspective resolution, children with TD do take into account the experimenter's viewpoint as a contextual resolution option. Children with ASD, on the other hand, were not affected by experimenter perspective, and lacked a clear contrast between the opposite deictic terms. They also had significantly higher error rates than children with TD, suggesting general difficulties in grasping the relational nature of the two deictic terms. This is in noteworthy contrast to another experiment (not reported here) on proximal and distal spatial terms (e.g., *this*, *that*) with the same participants, where children with ASD' deictic interpretations reflected very few struggles of such kind and were more aligned with children with TD's. The particular challenge of the complexity of perspectival options for spatial deixis, also witnessed in the strongly varying adult behavior, thus seems to cause difficulties for children with ASD in the present study, making spatial deixis resolution a rich area for future experimental research, both in general and with regards to population-specific challenges.

References

Bartak, L. and M. Rutter (1974). The use of personal pronouns by autistic children. *Journal of Autism and Childhood Schizophrenia 4*, 127–145.

Bates, D., M. Mächler, B. Bolker, and S. Walker (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software 67*(1), 1–48.

Burgess, N., K. J. Jeffery, and J. O'Keefe (1999). *The hippocampal and parietal foundations of spatial cognition*. Oxford: Oxford University Press.

Bühler, K. (1934/2011). *Theory of Language – The Representational Function of Language [Sprachtheorie]*. Amsterdam: John Benjamins.

Charney, R. (1980). Speech roles and the development of personal pronouns. *Journal of Child Language 7*, 509–528.

Chiat, S. (1982). If i were you and you were me: the analysis of pronouns in a pronoun-reversing child. *Journal of Child Language 9*, 359–379.

Clark, E. V. (1978). From gesture to word: on the natural history of deixis in language acquisition. In J. S. Bruner and A. Garton (Eds.), *Human Growth and Development*. Oxford: Oxford University Press.

Diessel, H. (2014). Demonstratives, frames of reference, and semantic universals of space. *Language and Linguistics Compass 8*(3), 116–132.

Diessel, H. and K. R. Coventry (2020). Demonstratives in spatial language and social interaction: An interdisciplinary review. *Frontiers in Psychology 11*, 133–146.

Fay, W. H. (1979). Personal pronouns and the autistic child. *Journal of Autism and Developmental Disorders 9*, 247–60.

Fillmore, C. J. (1971). Towards a theory of deixis. In *The PCCLLU Papers*. Department of Linguistics, University of Hawaii.

Fillmore, C. J. (1975). *Santa Cruz lectures on deixis, 1971*. Bloomington, IN: Indiana University Linguistics Club.

Fillmore, C. J. (1997). *Lectures on deixis*. Stanford: CSLI Publications.

González-Peña, P., M. J. Doherty, and P. Guijarro-Fuentes (2020). Acquisition of demonstratives in english and spanish. *Frontiers in Psychology 11*, 33–45.

Hobson, R. P., R. García-Pérez, and A. Lee (2010). Personcentred (deictic) expressions in autism. *Journal of Autism and Developmental Disorders 40*, 403–415.

Johnston, J. R. (1984). Acquisition of locative meanings: behind and in front of. *Journal of Child Language 11*(2), 407–422.

Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.

Lord, C. and R. Paul (1997). Language and communication in autism. In D. Cohen and F. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders*. New York: Wiley & Sons.

Loveland, K. A. (1984). Learning about points of view: Spatial perspective and the acquisition of i/you. *Journal of Child Language 11*, 535–556.

Loveland, K. A. and S. H. Landry (1986). Joint attention and language in autism and developmental language delay. *Journal of Autism and Developmental Disorders 16*, 335–349.

Lyons, J. (1977). *Semantics*, Volume 2. Cambridge: Cambridge University Press.

Oshima, D. Y. (2006). GO and COME revisited: What serves as a reference point? In *Annual Meeting of the Berkeley Linguistics Society 32*. The Berkeley Linguistics Society and The Linguistic Society of America.

Paillard, J. (Ed.) (1991). *Brain and space*. Oxford: Oxford Science Publications.

Piaget, J. (1928). *Judgement and reasoning in the child*. London: Routledge & Kegan Paul.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. `https://www.R-project.org/`.

Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology 41*, 1–48.

Raven, J., J. C. Raven, and J. H. Court (2000). *Raven manual: Section 3, standard progressive matrices, including the parallel and plus versions, 2000 edition*. Oxford: Oxford Psychologists Press.

Rocca, R., M. Wallentin, C. Vesper, and K. Tylén (2019). This is for you: Social modulations of proximal vs. distal space in collaborative interaction. *Scientific Reports 9*(14967), 1–14.

Shusterman, A. and P. Li (2016). Frames of reference in spatial language acquisition. *Cognitive Psychology 88*, 115–161.

Tager-Flusberg, H. (1999). A psychological approach to understanding the social and language impairments in autism. *International Review of Psychiatry 11*, 325–334.

Talmy, L. (1975). Semantics and syntax of motion. In J. P. Kimball (Ed.), *Syntax and Semantics*, Volume 4. New York: Academic Press.

Tosi, A., M. J. Pickering, and H. P. Branigan (2020). Speakers' use of agency and visual context in spatial descriptions. *Cognition 194*(104070), 124–129.

Tversky, B. and B. M. Hard (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition 110*(1), 124–129.

Wilkins, D. P. and D. Hill (1995). When GO means COME: Questioning the basicness of basic motion verbs. *Cognitive Linguistics 6*, 209–259.

Wilkinson, K. (1998). Profiles of language and communication skills in autism. *Mental Retardation and Developmental Disabilities Research Reviews 4*, 73–79.