

# ***This and that* back in context: Grounding demonstrative reference in manual and social affordances**

**Roberta Rocca** ([roberta.rocca@cc.au.dk](mailto:roberta.rocca@cc.au.dk))

Department of Linguistics, Cognitive Science and Semiotics,  
Aarhus University, Jens Chr. Skous Vej 2, 8000 – Aarhus C, Denmark

**Mikkel Wallentin** ([mikkel@cc.au.dk](mailto:mikkel@cc.au.dk))

Department of Linguistics, Cognitive Science and Semiotics & Center of Functionally Integrative Neuroscience,  
Aarhus University, Jens Chr. Skous Vej 2, 8000 – Aarhus C, Denmark

**Cordula Vesper** ([cvesper@cc.au.dk](mailto:cvesper@cc.au.dk))

Department of Linguistics, Cognitive Science and Semiotics & Interacting Minds Centre,  
Aarhus University, Jens Chr. Skous Vej 2, 8000 – Aarhus C, Denmark

**Kristian Tylén** ([kristian@cc.au.dk](mailto:kristian@cc.au.dk))

Department of Linguistics, Cognitive Science and Semiotics & Interacting Minds Centre,  
Aarhus University, Jens Chr. Skous Vej 2, 8000 – Aarhus C, Denmark

## **Abstract**

Spatial demonstratives, i.e. words like *this* and *that*, serve as important tools to establish joint attention, allowing interlocutors to flexibly share spatial reference schemes. However, little experimental work has investigated which perceptual and social factors drive speakers' choices of demonstrative forms. We used a novel experimental paradigm to explore 1) the role of relative placement of competing referents on the sagittal and lateral planes, 2) whether and how the presence of an addressee modulates the speaker's choice of demonstrative forms. We found that the choice of demonstratives is affected by the relative position of competing referents both on the sagittal and lateral plane. Furthermore, we found that the presence of an interlocutor shifts attraction for proximal demonstratives towards the shared space of reference, but only in collaborative contexts. Together, these results suggest that spatial deixis is grounded in a contrastive organization of space tightly coupled to manual and social affordances.

**Keywords:** demonstratives; social cognition; spatial cognition; spatial deixis

## **Introduction**

The ability to establish joint attention on objects or locations is a fundamental building block of human sociality (Tomasello, 2005). A wide spectrum of everyday activities relies on the ability to coordinate on and navigate joint attentional scenes. Natural languages are endowed with a large inventory of strategies that can be used for spatial referencing and coordination purposes (Tylén et al., 2010). Among them, spatial demonstratives, i.e. words like *this* and *that*, stand out, as attentional alignment is *integral* to their use (Diessel 1999). Demonstratives are prominent items in linguistic interaction. They are among the first lexical items to be mastered during development, and alongside

communicative pointing, they play a crucial role in bootstrapping language acquisition (Diessel, 1999; Diessel, 2006). Additionally, demonstratives can be regarded as cross-linguistic universals (Diessel, 2006).

In contrast to other strategies for verbal referencing, such as the use of nouns and descriptions, demonstratives carry minimal semantic specification of the intended referent, causing their interpretation to crucially hinge on the context of the utterance (Diessel, 1999; Levinson, 1983).

## **Physical context**

When used to refer to entities in the physical context of the utterance, demonstratives are typically coupled with visual signals such as pointing gestures (Clark, 1996; Cooperrider, 2016) or gaze cues (Perea-García et al., 2017), which deliver crucial information on the location of the intended referents relative to the speaker. In an EEG/ERPs experiment, Stevens & Zhang (2012) reported N400 effects for incongruence between demonstratives and object location only when the speaker and addressee in the scene established joint gaze on the referent. As demonstratives are used as attention aligning devices, participants perceived the absence of shared gaze as a violation in their use.

With the exception of very few languages, all demonstrative systems explicitly encode at least a minimal dyadic contrast between *proximal* and *distal* referents. More complex demonstrative systems display either more fine-grained distance-based contrasts, e.g. via explicit lexical encoding of medial distances from the speaker, or additional person-oriented contrasts, providing, for example, some specification on the position of referents relative to the addressee (Diessel, 1999).

Previous studies have experimentally investigated the motivations for this distance-based distinction, and provided empirical evidence for a mapping between the proximal/distal contrast in demonstrative systems and a functional

representation of space in body-centered coordinates. In a series of studies relying on a paradigm labelled the *memory game*, Coventry and colleagues asked participants to point at referents located at varying distances from the speaker on the sagittal axis and to refer to them by either a proximal or a distal demonstrative (Coventry et al., 2008; Coventry et al., 2014; Gudde et al., 2016). They established a mapping between distance-based contrasts in demonstratives and the distinction between peripersonal and extrapersonal space, consistent across a variety of genetically heterogeneous languages.

Interestingly, the choice of demonstratives along the proximal / distal axis inherits the characteristic flexibility of the boundary between peripersonal and extrapersonal space (Coventry et al., 2014). It has been shown that the use of proximal demonstratives is sensitive to manipulations of the scope of peripersonal space achieved with tool use (Longo & Lourenco, 2006). Moreover, just like physical distance and object attributes such as graspability and affective valence interact in perceptual judgements of reachability (Valdés-Conroy et al., 2012), the choice of demonstratives in dyadic systems is affected by perceptual parameters of the referent (e.g. visibility), as well as by psychological parameters such as ownership and familiarity (Coventry et al., 2014).

However, spatial demonstratives resist a rigid mapping onto body-centered representations of physical space. Bonfiglioli et al. (2009) exploited well-established interference effects between word meaning and movement planning and execution (Glover et al., 2004) to explore the semantics of spatial deixis. Participants were primed with either a proximal demonstrative or a distal demonstrative, then performed reach-for-grasp movements for objects located at two different distances in peripersonal space. They found that incongruence between demonstratives and spatial locations affected reaction times in movement execution. Together, these studies suggest that the use of demonstratives is not alone determined by physical distance between a single referent and the speaker, but rather reflects a flexible and context-sensitive implementation of the contrastive potential of the demonstrative system as a whole (Kemmerer, 1999; 2006).

### **Social context**

In a recent series of EEG studies, Peeters and colleagues (2015) have questioned purely egocentric proximity-based accounts of spatial deixis and stressed the role of the social context of utterance. Based on N400 effects, they reported a preference for proximal demonstratives for objects located in space shared between two interlocutors, irrespective of the distance of the referent from the speaker. Based on these results, the authors argue in favor of a reference frame

centered on the conversational dyad, rather than on the speaker alone. This proposal is in line with accounts of linguistic reference as a collaborative process (Clark, 1996). As pointed out within such frameworks, speakers design their communicative acts by actively taking into account the addressee's perspective and their common ground (Clark et al., 1983; Clark and Bangertner 2004).

### **The present study**

The aim of the present study was to address a number of outstanding questions related to how physical and social context influence demonstrative use. First, as previous experiments simplified reference resolution to single-referent contexts (which would not capture naturalistic situations of demonstrative use), we aimed to test how the presence of competing referents modulates the choice of demonstrative forms in interaction. *We hypothesized that proximal demonstratives are more likely to be used for referents relatively closer to the speaker on the sagittal axis.*

Secondly, existing paradigms have put the emphasis exclusively on the specifics of spatial deixis along the sagittal axis. It is, however, widely established that biomechanical constraints, such as handedness, may be a prominent source of asymmetries in perceptual space. Perceptual representations of peripersonal space underlying planning and execution of reaching movements rely on dynamic transformations between hand-centered and retinocentric coordinates (Makin et al., 2012). In the light of the hypothesis that the proximal/distal contrast is grounded on functional representations of space for reach and grasp (Coventry et al., 2014), a hand-centered frame of reference might indeed be crucial for the understanding of demonstrative reference. If the hand, rather than the locus of foveal fixation or the head, is the center of the deictic frame of reference, a lateralized bias for proximal demonstratives towards the hand used for pointing would be expected which has not been captured by previous experimental paradigms. Given these previous findings, and with all our participants being right-handed, *we hypothesized a right-lateralized bias in favor of proximal demonstratives.*

Last, existing literature has tackled deixis in individual contexts, without the presence of a conversational partner. However, while some languages lexicalize distinctions between locations of referents relative to an addressee, demonstrative choice in languages lacking an explicit encoding of the addressee's standpoint may still be affected by perceptual common ground between interlocutors (Peeters et al., 2015). *Therefore, we hypothesized that physical biases would be attenuated by the presence of a social partner in contexts of collaborative interaction.*

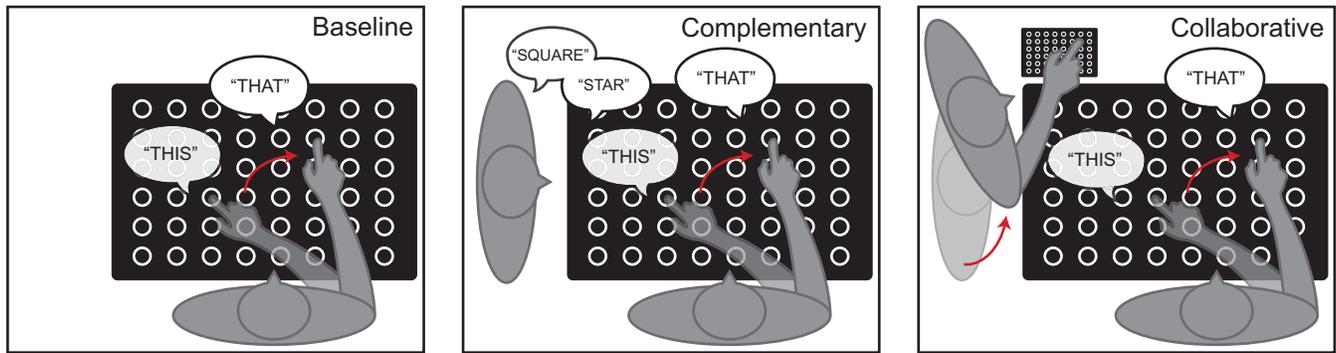


Figure 1: Experimental setup of the three conditions

To test our hypotheses, participants in the present study saw pairs of targets briefly appearing on a 40'' monitor lying horizontally. Their task was to point at the location of the two targets while referring to each of them by either a Danish proximal (“den her”) or a distal demonstrative (“den der”). Participants were tested under three conditions: an individual baseline, a complementary condition, where a confederate performed a task unrelated to the participant’s pointing, and a collaborative condition, in which the participant’s pointing was functional to the confederate’s task. The hypotheses were preregistered on the Open Science Framework prior to conducting the experiment. The preregistration is available here: [osf.io/gjnf9](https://osf.io/gjnf9).

## Methods

### Participants

80 right-handed participants (female = 43, age range = 19–48, median = 26,  $sd = 7.6$ ) with Danish as first language took part in the experiment at the Cognition and Behavior Lab at Aarhus University in return for a monetary compensation. The authors and two student assistants took turns in the role of experimenter (live-coding the subject’s responses) and confederate.

### Design & Procedure

The task was presented as a spatial working memory test (in line with Coventry et al., 2008; 2014). In each trial, a grid of circles would appear on the screen for 500 ms. Then the grid disappeared and two target shapes (circles, triangles, squares, hexagons, stars) appeared on the screen for a random interval between 200–800 ms. Then the grid would reappear and the participants were prompted to designate the target positions. The position of targets was randomized across trials. Subjects were instructed to remember the locations of targets, then point at them while referring to the objects with the Danish demonstratives “den her” or “den der”. They were explicitly instructed to use both demonstrative forms in each trial, and were reminded to do so whenever they disregarded the rule. No explicit instructions were given on the order of

the points nor on the order of deictic forms. There were 132 trials per condition per subject.

Participants performed the task across three conditions. In the *baseline condition*, subjects performed the task alone. In the *complementary condition*, a confederate stood to the left of the subject and named the target shapes (e.g. “star, circle”) after the participant was done pointing. There was no interaction between the two tasks, and therefore neither the participant nor the confederate depended on the information provided by the other to perform their own task. In the *collaborative condition*, the confederate would close her eyes during target exposure and only opened them after a click sound. The participant then pointed at the location of both targets so to allow the confederate to report them on a touch screen device placed next to the big screen. The baseline was always performed first. The order of the complementary and collaborative conditions was counterbalanced.

### Analysis

The relative distances between the x coordinates and the y coordinates of the two targets were used as predictors for a mixed effects logistic regression using the *glmer* function from *lme4* package in RStudio. For each trial, one of the two targets (henceforth: T1) was randomly selected and logged as target of interest. The relative distances on each of the axes were computed by subtracting the x coordinates and the y coordinate of the competitor target (henceforth: T2) from those of T1. The relative distance on the x axis took positive values if T1 was further to the right than T2, whereas their distance on the y axis took positive values if T1 was further away from the speaker than T2 on a sagittal axis. The *fixed effects* structure of the model included the relative distance between the two targets on the y axis, the relative distance between the two targets on the x axis, and condition, as well as all interactions.

The demonstrative form (proximal or distal) chosen to refer to T1 was used as outcome variable in the model. The distal demonstrative (“den der”) was set as reference level, while the proximal form (“den her”) was coded as success outcome.

The *random effect* structure included random intercepts for each participant as well as random slopes for relative distance on the y axis.

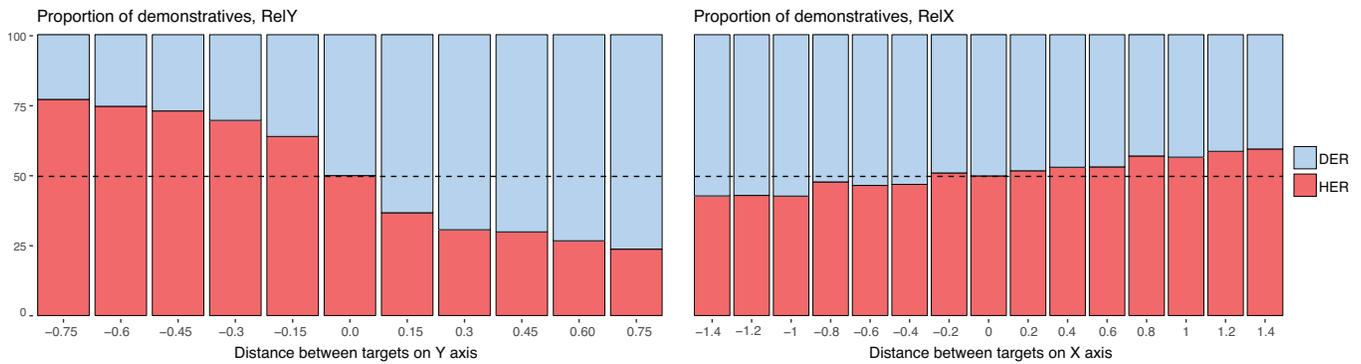


Figure 2: proportion of demonstratives across values of *RelY* (left) and *RelX* (right) across all conditions

Parameters were estimated using maximum likelihood estimation with Laplace approximation. The power simulation for the model is reported in the preregistration, yielding 70-100% power for fixed effects and two-way interactions. Planned contrasts for the categorical predictor compared the subject's behavior in the baseline condition with cumulative behavior in the social condition, as well as the complementary condition against the collaborative condition.

## Results

After discarding invalid responses, a total of 31394 out of 31680 data points was included in the analysis. As shown in Figure 2 (left), the proportion of proximal demonstratives decreases as a function of increases in the value of *RelY*, i.e. as T1 moves further away from the speaker on the sagittal axis relative to T2. Figure 2 (right) displays the pattern for *RelX*: there was an increase in proportion of proximal demonstratives as a function of an increase in the value of *RelX*, i.e. as T1 moves further to the right relative to T2.

A mixed effects logistic regression model with *RelX*, *RelY* and *condition* as predictors, and including all interactions, confirms the statistical reliability of these patterns. The model displays a significant effect of *RelY*,  $\beta = -2.59$ ,  $se = 0.27$ ,  $z = -9.7$ ,  $p < .001$  and of *RelX*,  $\beta = 0.32$ ,  $se = 0.02$ ,  $z = 16.77$ ,  $p < .001$ .

Planned contrasts reveal a significant interaction between *RelX* and *Condition* when comparing the *complementary* and *collaborative* condition,  $\beta = 0.05$ ,  $se = 0.02$ ,  $z = 2.17$ ,  $p < .05$ . No such effect is observed when cumulatively comparing the baseline to the social conditions,  $\beta = -0.001$ ,  $se = 0.01$ ,  $z = -0.079$ ,  $p = .93$ .

Moreover, the interaction between *RelY* and *Condition* reaches statistical significance both in the contrast between baseline and the two social conditions,  $\beta = -0.07$ ,  $se = 0.03$ ,  $z = -2.51$ ,  $p < 0.01$ , and between the *complementary* and *collaborative* condition,  $\beta = -0.11$ ,  $se = 0.05$ ,  $z = -2.45$ ,  $p < .01$ . None of the three-way interactions reached statistical significance.

## Discussion

In the following, we discuss the results with respect to our three hypotheses.

**Relative distance on the sagittal axis** As shown by the main effect of *RelY*, the relative distance between targets on the sagittal axis had an impact on the use of demonstratives. As the sagittal distance between the two referents increased, the likelihood of observing a distal demonstrative progressively increased. While the latter point is in line with the well-established preference for distal demonstratives for referents placed far from the speaker (Coventry et al., 2008), such a finding adds to current knowledge along some intriguing dimensions. First, it provides empirical evidence for the hypothesis that demonstrative reference is grounded on the construction of a contrastive space of competing referents, in addition to a mapping between deictic space and near/far space. In other words, in the context of competing referents, the organization of deictic space is set up contrastively, by taking into account the interplay between distances of the intended referent from the speaker's body and from competing referents.

**Right-lateralized bias within peripersonal space** Across all conditions, we found a right-lateralized preference for proximal demonstratives. The preference for proximal demonstratives was strengthened as the distance between the target referent and the competing referent increased. This finding is in line with previous research hypothesizing a lateralized bias in peripersonal space, and can be attributed to space being encoded in hand-centered coordinates, which facilitates fast execution of hand movements by reducing the load of sensorimotor transformations (Makin et al., 2012). This correspondence between the organization of demonstrative space and hand-based encoding of space is in line with the link between the proximal/distal distinction and reach for grasp actions, which has previously been established in the literature on demonstratives use (Coventry et al., 2014; Bonfiglioli et al., 2009). While Coventry and colleagues (2008; 2014) frame this relationship in terms of binary (though flexible) distinction between space within-reach and out of reach, Bonfiglioli and coworkers (2009) argue that the conceptual space for *this* and *that* is grounded on differences in distance even within peripersonal space.

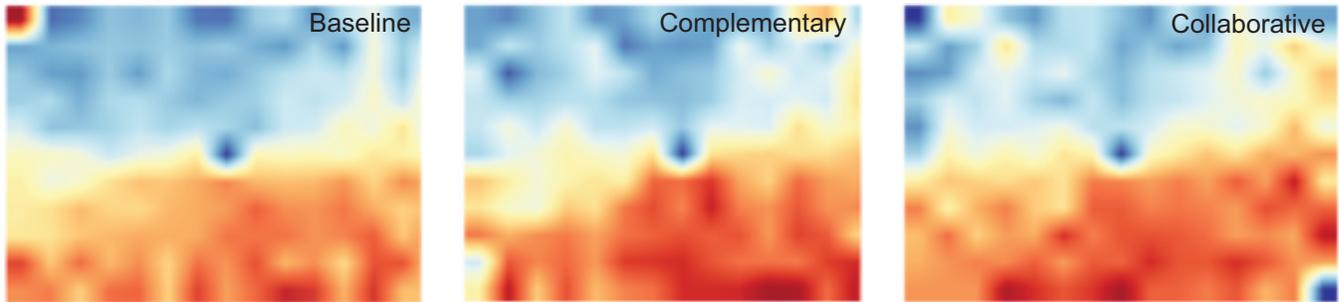


Figure 3: Heat maps displaying the proportion of proximal demonstratives across values of  $RelX$  (x-axis) and  $RelY$  (y-axis)

Together, our findings provide evidence in favor of proximal/distal distinctions being grounded on gradient affordance for manual interaction rather than on an absolute mapping. While the present investigation does not fully explore manipulations of target locations going significantly beyond peripersonal space, the paradigm allowed to show that the proximal/distal organization of space is indeed grounded in biomechanics, with objects more readily affording grasp being more likely to be labelled via a proximal demonstrative.

This further underlines the integral role of multimodal components in deictic reference, as the right-lateralized bias in the organization of space probably reflects a bias towards the pointing hand. However, in the experiment we did not manipulate which hand is used for pointing nor did we include any left-handed participant. Therefore, the current results might either be due to a bias in favor of the hand used for pointing (independently of handedness) or in favor of the dominant hand (independently of the hand used for pointing). Further research including relevant manipulations is required in order to disentangle such possibilities. An additional explanation is that asymmetries are due to a general perceptual bias. However, previous studies have mostly reported left-lateralized biases in perceptual tasks and visual imagery, which have been attributed to reading direction (Jewell & McCourt, 2000; Stoustrup & Wallentin, 2017). Our results point in the opposite direction, thus suggesting that biomechanical constraints tend to override purely perceptual sources of asymmetries when motor components are essential for the task at stake.

It is to be pointed out that we initially predicted lateralized biases to be detected in the form of an interaction between  $RelX$  and  $RelY$ , as we expected it to apply *only* to referents with low absolute values of  $RelY$ . The observed main effect of  $RelX$  rather suggests that the right-lateralized bias is independent of the position of the targets on the y-axis, which shows that such bias might be even more prominent than expected in influencing demonstrative production. This consideration also applies to the interaction between  $RelX$  and *Condition* reported below, which was initially expected in the form of a three-way interaction with  $RelY$ .

**Social presence and collaboration** For demonstration purposes, Figure 3 displays the distribution of proximal demonstratives as a function of variation in  $RelX$  and  $RelY$ . The heat maps are oriented so to display higher values of  $RelX$  towards the right, and higher values of  $RelY$  towards the

top. As figure 3 suggests, the baseline and the complementary condition display a similar pattern, with a slightly more pronounced bias in the complementary condition. In the complementary condition, the confederate's task was to name which shapes lit up on the screen. When the confederate takes up the semantic part of the task, the speaker's increased focus on the spatial location of the referents resulted in a more pronounced right-lateralized bias.

However, the bias is significantly attenuated in the collaborative condition, i.e. when the speaker and addressee are engaged in a task involving actual collaboration and, therefore, functional, communicative pointing. This result provides evidence for the fact that, when pointing, the participant factors in the position of the addressee, which induces a shift in the proximal space towards the space shared between the two interlocutors. Interestingly, this is not the case for mere co-presence of another participant (in the complementary condition), but rather requires involvement of the interlocutors in a collaborative interaction. In this case, since the information conveyed by the speaker's pointing is functional to the addressee's task, the speaker might spontaneously facilitate the interlocutor's task by adjusting her proximal space towards shared space.

The effect of  $RelY$  also seems to be modulated across conditions, with the preference for proximal demonstratives for closer referents becoming weaker in the two social conditions compared to the baseline, an effect which is likely to be driven by the collaborative condition. Although it has to be acknowledged that this effect was not part of our initial set of hypotheses, it affords an interpretation compatible with our previous remarks. Indeed, the effect suggests that the purely hand-based contrastive space set up in the baseline is attenuated by the presence of an addressee in favor of an organization of space which takes social components into account.

Moreover, the heat map for the collaborative condition indicates higher frequency of proximal demonstratives for negative  $RelX$  and higher  $RelY$  (centre left of the map) than the complementary condition and the baseline, which corresponds to proportionally more proximal demonstratives for cases in which T1 is located more towards the left of the speaker (and therefore closer to the addressee) but further away from the speaker. This suggests that such effect might be mostly driven by data points with high values of  $RelX$ .

## Conclusion

We investigated the effect of object location and social presence in use of demonstratives. Our results suggest that participants' tendency to use proximal demonstratives increases as the target of interest gets closer to the speaker and further away from the competing target, which suggests that deictic space is organized as a contrastive space rather than relying uniquely on a fixed mapping between peripersonal and extrapersonal space.

Additionally, we observed a right-lateralized bias in the use of proximal demonstratives. This provides evidence towards a tight link between the organization of space in spatial deixis and action, under the hypothesis that proximal demonstratives are more likely to be used for referents affording easier manual interaction. Finally, we observed that both the effect of relative distance on the y axis and the effect of relative distance on the x axis are modulated by the presence of an addressee cooperating with the speaker in solving a shared task. Speakers shift their proximal space towards shared space, a finding which suggests that the organization of space in demonstrative reference is tightly coupled to social affordances.

## Acknowledgements

This project was funded via a seed-funding grant from the Interacting Minds Centre at Aarhus University. Roberta Rocca is funded by the DCOMM grant (EU H2020 ITN Marie Skłodowska-Curie Actions; grant agreement 676063). We wish to thank Sergio Gonzalez de la Higuera Rojo and Caroline Kildahl for their assistance with data collection.

## References

- Bonfiglioli, C., Finocchiaro, C., Gesierich, B., Rositani, F., & Vescovi, M. (2009). A kinematic approach to the conceptual representations of this and that. *Cognition*, *111*(2), 270-274.
- Clark, H. H. (1996). *Using language*. Cambridge University Press.
- Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground at the understanding of demonstrative reference. *Journal of verbal learning and verbal behavior*, *22*(2), 245-258.
- Clark, H. H., & Bangerter, A. (2004). Changing ideas about reference. In *Experimental pragmatics* (pp. 25-49). Palgrave Macmillan UK.
- Cooperrider, K. (2016). The co-organization of demonstratives and pointing gestures. *Discourse Processes*, *53*(8), 632-656.
- Coventry, K. R., Valdés, B., Castillo, A., & Guijarro-Fuentes, P. (2008). Language within your reach: Near–far perceptual space and spatial demonstratives. *Cognition*, *108*(3), 889-895.
- Coventry, K. R., Griffiths, D., and Hamilton, C. J. (2014). Spatial demonstratives and perceptual space: describing and remembering object location. *Cogn. Psychol.* *69*, 46–70.
- Diessel, H. (1999). *Demonstratives: Form, function and grammaticalization*, John Benjamins Publishing.
- Diessel, H. (2006). Demonstratives, joint attention, and the emergence of grammar. *Cognitive linguistics*, *17*(4), 463-489.
- Glover, S., Rosenbaum, D. A., Graham, J., & Dixon, P. (2004). Grasping the meaning of words. *Experimental Brain Research*, *154*(1), 103-108.
- Gudde, H. B., Coventry, K. R., & Engelhardt, P. E. (2016). Language and memory for object location. *Cognition*, *153*, 99-107.
- Jewell, G., & McCourt, M. E. (2000). Pseudoneglect: a review and meta-analysis of performance factors in line bisection tasks. *Neuropsychologia*, *38*(1), 93-110.
- Kemmerer, D. (1999). “Near” and “far” in language and perception. *Cognition*, *73*(1), 35-63.
- Kemmerer, D. (2006). The semantics of space: Integrating linguistic typology and cognitive neuroscience. *Neuropsychologia*, *44*(9), 1607-1621.
- Levinson, S. C. 1983. *Pragmatics*. Cambridge University Press.
- Longo, M. R., & Lourenco, S. F. (2006). On the nature of near space: Effects of tool use and the transition to far space. *Neuropsychologia*, *44*, 977-981.
- Makin, T. R., Holmes, N. P., Brozzoli, C., & Farnè, A. (2012). Keeping the world at hand: rapid visuomotor processing for hand–object interactions. *Experimental brain research*, *219*(4), 421-428.
- Perea-García, J. O., Ehlers, K. R., & Tylén, K. (2017). Bodily constraints contributing to multimodal referentiality in humans: The contribution of a de-pigmented sclera to proto-declaratives. *Language & Communication*, *54*, 73-81.
- Peeters, D., Hagoort, P., and Özyürek, A. (2015). Electrophysiological evidence for the role of shared space in online comprehension of spatial demonstratives. *Cognition* *136*, 64–84
- Stevens, J., & Zhang, Y. (2013). Relative distance and gaze in the use of entity-referring spatial demonstratives: An event-related potential study. *Journal of Neurolinguistics*, *26*(1), 31-45.
- Stroustrup, S., & Wallentin, M. (2017). Grammatical category influences lateralized imagery for sentences. *Language and Cognition*, 1-15.
- Tomasello, M., Carpenter, M., Call, J., Behne, T. and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences*, *28*, 675–91; discussion, 691-735.
- Tylén, K., Weed, E., Wallentin, M., Roepstorff, A., & Frith, C. D. (2010). Language as a tool for interacting minds. *Mind & Language*, *25*(1), 3-29.
- Valdés-Conroy, B., Román, F. J., Hinojosa, J. A., & Shorkey, S. P. (2012). So far so good: Emotion in the peripersonal/extrapersonal space. *PLoS one*, *7*(11), e49162.