Title: Spatial deixis back in context: the social space of demonstrative reference in interaction
Authors: Roberta Rocca[1], Mikkel Wallentin[1,2], Cordula Vesper[1], Kristian Tylén[1,3]
Affiliations: [1]School for Communication and Culture, Aarhus University; [2]Centre for Functionally Integrative Neuroscience, Aarhus University; [3]Interacting Minds Centre, Aarhus University

Short introduction

It is usually claimed that, at least in dyadic systems, the choice of so-called *proximal* and *distal* demonstratives encodes differences in *egocentric* distance between the referent and the speaker (e.g. Diessel, 1999, 2014; Coventry et al., 2008). However, recent studies have pointed at a more complex picture, with the relative distance between referents (Bonfiglioli et al., 2008) and social aspects of the interaction at stake (Peeters et al., 2015) playing a role in the choice of demonstrative forms.

In this experiment, we aim at introducing a new experimental paradigm to both replicate previous findings and provide original contributions to the debate.

We will focus specifically on the use of spatial demonstratives in Danish (i.e. *den her* and *den der*), with the purpose of:

- Replicating previous findings hinting at a significant role of relative distance between competing referents in demonstrative choice within peripersonal space;
- Investigate whether a bias can be observed in the choice of demonstrative forms on the lateral axis of space, with proximal forms being preferred for objects in the side of space ipsilateral to the pointing hand;
- Investigate whether the presence and the role of an interlocutor in a communicative interaction influences the speaker in her choices of either demonstrative form.

A. Hypotheses

1. Relative distance between competing referents

In line with previous research (Bonfiglioli et al., 2008), we hypothesize that, when competing referents are present within peripersonal space, the likelihood of using a proximal demonstrative for a referent increases as a function of the proximity of the target to the speaker, relative to the position of the competing target.

In the experiment, two targets (henceforth: T1 and T2) will light up on a screen, lying on a horizontal surface (more detail on the task is provided in the following sections). We will refer to the coordinate of each target on the sagittal axis as the *y coordinate* of the target. We will refer to the coordinate of each target on the lateral (left to right) axis as the *x coordinate* of the target. The relative distance between T1 and T2 on the sagittal axis will be computed as the difference between the y coordinate of T1 and the y coordinate of T2 (henceforth: *RelY*). Values for y coordinates increase as the targets are further from the speaker: therefore, the difference between y coordinates of T1 and T2 will be positive if is T1 is further away from the speaker than T2 (more detail on the coordinate system is provided in the *Methods* section, Figure 1). The outcome variable is a binary variable encoding whether or not a proximal demonstrative is chosen to refer to T1 (henceforth: *Demonstrative*).

This hypothesis will be tested by assessing the contribution of *RelY* in predicting the probability of using a proximal demonstrative, in a logistic regression model with *Demonstrative* as outcome variable. We expect to observe a robust negative effect of *RelY* on the probability of observing a proximal demonstrative.

2. Rightward bias for proximal demonstratives

Existing literature on the role of demonstratives has focused on investigating the mapping between proximal/distal demonstratives and near/far referents by manipulating uniquely the distance of referents from the speaker on a sagittal axis. However, it is known that asymmetries can be observed between left and right hemifield of peripersonal space both in the perceptual and in the motor domain. Demonstratives are consistently used in co-occurrence with a pointing gesture. Therefore, we hypothesize that asymmetries in motor routines transfer to the use of demonstrative forms, with the use of proximal demonstratives being more likely for referents ipsilateral to the dominant hand.

This hypothesis will be tested by assessing whether the relative distance of referents on the horizontal axis (henceforth: *RelX*) significantly modulates the effect of *RelY*. We will therefore test whether a significant interaction can be observed between the *RelY* and *RelX*, expecting a stronger increase in likelihood of using a proximal demonstrative for closer referents located further towards the right of the participant.

Relative distance on the x axis is measured as difference between the x coordinate of T1 and the x coordinate of T2. Positive values correspond to T1 being to the right of T2. Again, a more detailed representation on possible target positions and coordinate systems for targets is provided in the *Methods* section (see Figure 1).

3. Effect of presence and role of an interlocutor

Previous experimental paradigms on demonstratives have focused uniquely on deixis in abstract, non-interactive contexts. However, the use of demonstratives is highly dependent on properties of the communicative interactions in which they are embedded. For instance, it has been shown that the distinction between space shared by two interlocutors and non-shared space plays a crucial role in the choice of which demonstrative form to use (Peeters et al., 2015). Referents placed in the region of space shared by the interlocutors tend to be referred to with a *proximal* demonstrative form, whereas objects placed outside this region tend to be referred to with a *distal* form, regardless of them being placed within the peripersonal or the extrapersonal space of the speaker. The mapping between proximal/distal forms and the position of referents in space seems therefore to be tied to a conceptualization of space as both *metric* and *social*.

In order to shed light on how the presence of an interlocutor influences the speaker's choice of a proximal or a distal demonstrative form, the experiment includes a non-interactive baseline, where the subject performs the task individually, and two interactive conditions. In the first interactive condition (henceforth: *Complementary* condition), a confederate stands to the left of the subject and performs a task complementary to the subject's, i.e. a task which does not entail the use of information conveyed by the subject's referential act. In the second interactive condition (henceforth: *Collaborative* condition), a confederate stands to the left of the subject and collaborates with the subject in performing a joint task. In the joint task, the subject points at the location of the two targets *for* the confederate, who taps the corresponding target locations on a grid displayed on a touch screen device.

We expect the presence of an interlocutor to modulate the pattern of use of demonstrative forms. More specifically, we expect the rightward bias for proximal referents (predicted in the previous paragraph) to be significantly attenuated. Under this hypothesis, compared to the baseline, for the complementary condition we would observe a higher likelihood of using proximal forms for targets closer to the speaker on the y axis and placed closer to/within the region of space shared between subject and confederate, namely the surface of the screen towards the left of the speaker. We expect this to be consistent both when comparing the baseline condition with the complementary condition, and when comparing the baseline condition with the collaborative condition, with a stronger effect in the latter case.

We expect to observe this in terms of significant three-way interactions between *RelY*, *RelX* and *Condition*.

B. Methods

B.1 Independent variables

B.1.1 Condition: manipulation of presence and role of an interlocutor in the interaction. Participants are asked to perform the task across three conditions:

- Baseline: subjects perform the task alone. No potential interlocutor is present in their surroundings;
- Complementary: a confederate stands to the left of the subject and names the objects lighting up on the screen after the subject has referred to them via demonstratives. No interference between the two tasks, and therefore neither the participant nor the confederate depend on the information provided by the other for solving their own task;
- Collaborative: a confederate does not have access to the targets and thus depend on the participant to indicate the location of the two targets to solve her task. Given the collaborative nature of the interaction, the subject's referential act is endowed with communicative intention.

It is a fully within-subject manipulation.
While the baseline is always performed at the beginning, the order to the complementary and collaborative condition is counterbalanced across participants.

B.1.2 Position of targets

Different shapes appear as targets on the screen (circles, triangles, squares, hexagons, stars). Two target shapes light up simultaneously at each trial. The position of targets is manipulated across trials. Targets can appear at two of the 48 potential locations exemplified in Figure 1 (below). Pairs of randomly chosen x,y combinations are selected at each trial for both target objects.
The figure below illustrates the potential positions of targets on the screen, as well as the scale and orientation of the coordinate system.

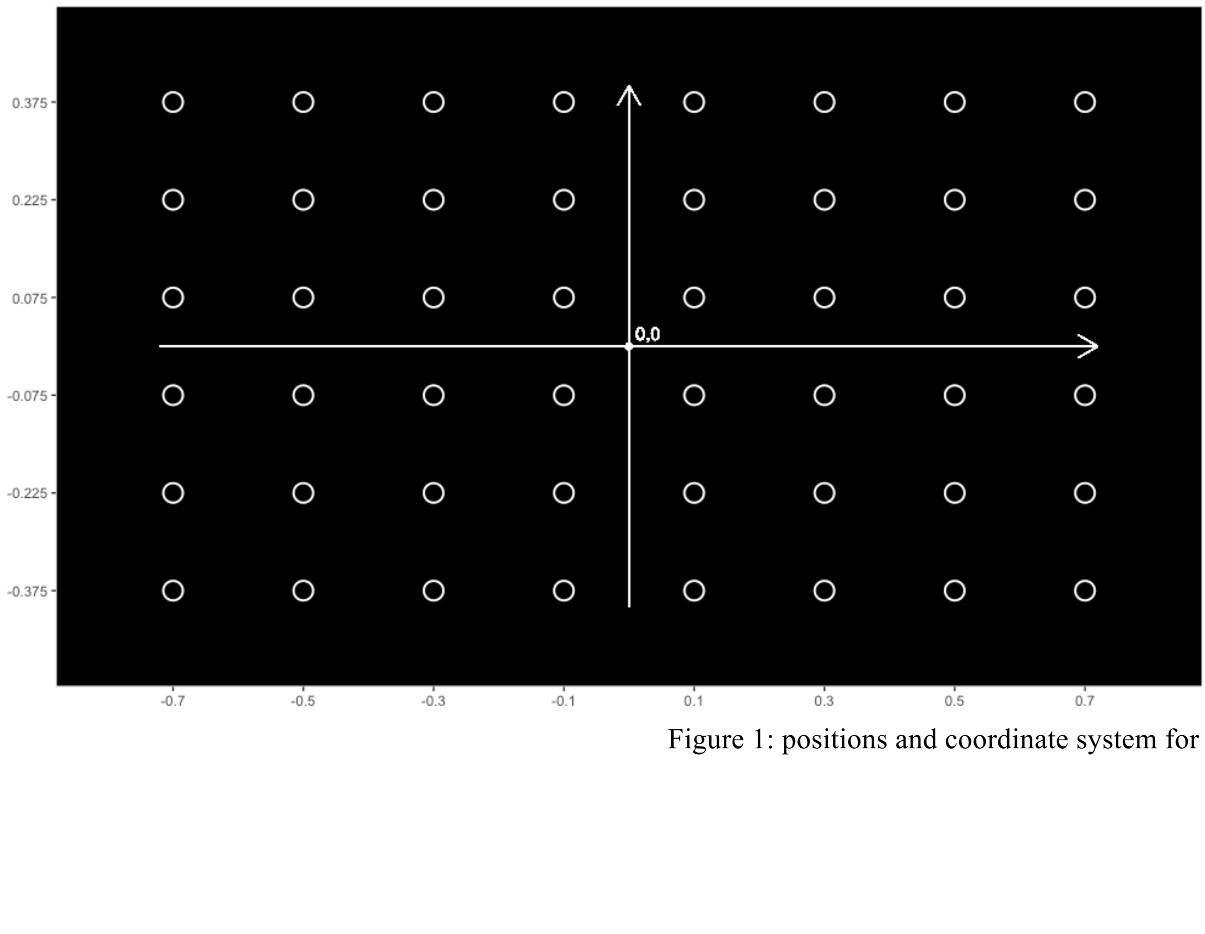

Figure 1: positions and coordinate system for targets

There are 132 trials divided into two blocks per condition per subject.
Coordinates on the x and y axis are centred on 0, with (0,0) corresponding to the centre of the screen. The x axis is oriented so to have positive values to the right of the screen, and negative values to the left. The y axis is oriented so to have positive values towards the top of the screen (further from the speaker), and negative values towards the bottom (closer to the speaker). Values are bounded between -0.7 and 0.7 on the x axis, between -0.375 and 0.375 on the y axis.

In our logistic regression model, absolute locations of the targets will not be modelled. In order to test for the effect of relative location of competing targets, we will model:

- Relative distance between T1 and T2 on x axis, *RelX*.
- Relative distance between T1 and T2 on the y axis, *RelY*.

All combinations of relative target locations are crossed across levels of the *Condition* variable.
Detailed explanation on variables of interest and predicted main effects and interactions are provided in the *Hypotheses* section.

***B.2* Dependent variables**
The outcome variable is a binary variable coding for deictic form used to refer to T1. 'Den her' is the Danish proximal demonstrative, whereas 'Den der' is the distal form. 'Den her' is coded as *success* outcome, i.e. the regression models the probability of observing 'Den her' as a function of *RelY*, *RelX*, *Condition* and their relevant interactions.

***B.3* Model**
Data will be analysed using the *glmer* function from *lme4* package in the statistical software R.
We will assess the contribution of regressors within the model specified as follows:

```
model <- glmer(Demonstrative ~ RelY + RelX + RelY:RelX + RelY:RelX:Condition + (1+ RelY|Subj_ID), family = binomial(logit))
```

Parameters will be estimated using maximum likelihood estimation with Laplace approximation.
RelX is included as a covariate, but not tested for an effect of interest.
Inferences will be drawn using Wald test, default in:

```
summary(model)
```

In order to further assess the contribution of each predictor to the model, comparisons with a null random-effects only model and with nested models might be conducted.

***B.4* Power Analysis**
We conducted a power analysis via simulation. The code used for simulations is attached to this draft. The simulation assumed a .6 standard deviation in both subjects' intercepts and slopes and a correlation of .2 between the two. We generated data with the logistic link function assuming a value of 0 for the intercept, and both 0.5 and 0.7 effect sizes for each of the effects of interest.
In the simulation, power was computed on the basis of the percentage of successes in detecting a significant effect of the same sign as the parameter of the generative distribution.
Detail on the result of the simulation is attached. Simulating 80 participants for 3 conditions and 132 trials per condition yields, for effect size = .5:
- 100% power for the main effect of RelY;
- 84% power for the interaction RelY:RelX;

- 56% power for the interaction RelY:RelX:Condition_Complementary;
- 55% power for the interaction RelY:RelX:Condition_Collaborative

For effect size = .7, the simulation yields:
- 100% power for the main effect of RelY;
- 98% power for the interaction RelY:RelX;
- 87% power for the interaction RelY:RelX:Condition_Complementary;
- 81% power for the interaction RelY:RelX:Condition_Collaborative

B.5 Planned Sample

80 right-handed participants with Danish as first language and age ranging between 18 and 40 will be recruited for the study.

Participants will be recruited via the SONA pool at COBE lab (Cognition and Behavior lab), Aarhus University. Data will be collected in the premises of COBE lab, by the authors and a student assistant hired for the project. A Danish native speaker will be hired as confederate for the collaborative and complementary condition.

Data collection will be terminated once the planned sample size has been reached.

Data from participants consistently not acting in accordance with the given instructions will be discarded. Relevant instructions include using both demonstrative forms in each trial, waiting until the given auditory signal before responding, bringing the pointing arm and hand back to the starting position after each trial.

B.6 Experimental setup

The experimental setup includes:

- A 40'' flat screen lying horizontally on a stable surface;
- A webcam, placed on a tripod at a height of around 1m above the screen, recording both the screen and the participant all throughout the experiment.
- A microphone
- A booth separated from the participant via a vertical surface which prevents visibility both ways. Participant responses are still clearly audible from the booth, which is essential for the live coding (see *Procedure* section).
- A laptop computer connected to the webcam and the microphone, placed inside a booth invisible to the participant, where the video is streamed during the experiment in order to allow the experimenter to monitor the subject's behaviour and responses. The video (including audio) is recorded on this laptop.
- A laptop computer connected to the screen, delivering the stimuli and used to live code the subject's responses based on the video streaming.
- A tablet/touch screen device, used only in the collaborative task, to allow the confederate to code target locations based on the subject's response. The touch screen device displays a grid isomorphic to the one on the screen, where the confederate touches the corresponding target location.

B.7 Procedure

Participants are introduced to the experiment, sign a consent form.

If necessary, adjustments are made to the height of the surface where the screen lies, and set at a distance optimal for performing pointing gestures. The webcam and microphone are placed so to record the entire area of interest and to capture audio information optimally.

Participants are then provided with detailed instructions on the task. Subjects are blinded to the purpose of the experiment. The task is presented as meant to test short term memory for object location. They are instructed to pay attention to the shapes lighting up on the screen, remember the exact locations and point at them by simultaneously referring to the objects with "Den her" (= "this") and "Den der" (= "that") after the targets disappear.
They are explicitly instructed to use both demonstrative forms in each trial, and are reminded to do so during the experiment if they disregard the rule.
They are instructed to point and refer to the objects only after they disappear, which is signalled by a short beep tone.

In each trial, a grid with possible referent locations appears for 500 ms. Afterwards, the grid disappears, while two shapes light up in two of the potential locations from the grid. Targets disappear after a variable delay, replaced again by the initial grid. The appearance of the grid is coupled with a beep sound, to signal that the participant is now allowed to respond. Participants point at target locations one by one, with each point being accompanied by either the verbal response '*den her'* or '*den der'*.
Trials are considered valid only when both forms are used within a trial.
There is no explicit instruction on the order of the points nor in the order of use of deictic forms.
The experimenter codes the responses live from the booth, and triggers manually the next trial.
There are three blocks of 66 trials each per condition.
The experimental trials are preceded by 3 practice trials, where the subject becomes accustomed with the procedure.

In the *Complementary* condition, the confederate stands to the left of the participant.
Before the practice trials, the list of shapes is displayed on the screen in order to provide the confederate with the appropriate lexicon to perform the task.
In this condition, the experiment follows the same procedure as the *Baseline* condition. The only major difference is that in each trial, after the subject has pointed at *both* targets, the confederate names the two shapes (e.g. "star, circle"). Live coding is performed by the experimenter following the same procedure as the baseline condition.

In the *Collaborative* condition, the confederate stands on the left of the participant, holding a touch screen device in his hand. The confederate turns away from the screen while the targets are displayed, then turning back towards the screen after the beep sound. After the subject has pointed at both targets, the confederate reports the pointed targets by touching the relevant locations on the touch screen device.

Conditions are presented in sequential order, with the baseline condition first, being followed by complementary and collaborative condition. Complementary and collaborative conditions are counterbalanced across participants.

At the end of the experiment, participants are debriefed and asked to sign the post-experiment consent form. The confederate is blind to the purpose of the experiment.
Video recordings are kept to check for values missing from the live coding.

C. ANALYSIS PLAN

The analysis is conducted in the form of a mixed effects logistic regression.

The *fixed effects* structure includes:

- *RelY*, i.e. the relative distance between the two targets on the y axis. It is computed as the difference between the y coordinate for T1 and the y coordinate for T2, therefore resulting in a positive value if T1 is further from the speaker than T2, and in a negative value if T2 is further than T1. Included as *independent variable*, under the hypothesis that closer objects are more consistently associated with proximal demonstratives.
- *RelX*, i.e. the relative distance between the two targets on the x axis. It is computed as the difference between the x coordinate for T1 and the x coordinate for T2, therefore resulting in a positive value if T1 is more towards the right than T2, and in a negative value if T2 is more towards the right. Included as *covariate*.
- The interaction between *RelX* and *RelY*. Introduced to test the hypothesis that the position of the referent on the X axis modulates the effect of its relative position on the y axis on the outcome variable;
- The interaction between *RelX*, *RelY* and *Condition* (discrete variable, having the *baseline* condition as reference level). The variable condition is hypothesized to be a *moderator* variable. Tests the hypothesis that the spatial map for demonstrative use is modulated by the presence and role of an interlocutor, with higher likelihood of using the proximal demonstrative for closer targets more towards the left of the subject in both the interactive conditions, compared to the baseline.

The *random effect* structure includes random intercepts for each subject as well as random slopes for *RelY* (motivated by variability in arm length, which might correspond to high variability in the structure of peripersonal space).
The outcome variable, labelled as *Demonstrative*, is a binary value coding for the demonstrative used to refer to. Which target is T1 and which one is T2 is determined randomly.
Detailed information on the relevance of each predictor for the hypotheses is provided in the *Hypotheses* section. Information on parameter estimation and inference is provided in the *Model* subsection of *Methods*.

p-values will be corrected for multiple comparisons using *Bonferroni correction*.
Missing values will be handled via *listwise deletion*.

D. DATA COLLECTION

Only complete pilot data from 2 participants have been collected. Data collection from the full sample has not begun yet. It is planned to take place from ultimo October 2017 to December 2017.

E. LITERATURE

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.

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. doi: 10.1016/j.cogpsych.2013.12.001

Diessel, H. (2006). Demonstratives, joint attention, and the emergence of grammar. *Cognitive linguistics*, *17*(4), 463-489.

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. doi: 10.1016/j.cognition.2014.10.010