Spatial Demonstratives and Physical Control

Angelo Mattia Gervasi (angelomattia.gervasi@uniroma1.it)
Sapienza University of Rome
Via degli Apuli 1, 00185, Rome, Italy
UCBL1 University of Lyon
Lyon Neuroscience Research Centre, INSERM U1028
Lyon, France

Anna M. Borghi (anna.borghi@uniroma1.it)
Sapienza University of Rome
Via degli Apuli 1, 00185, Rome, Italy
Institute of Cognitive Sciences and Technologies, CNR
Via Romagnosi 18A, 00196, Rome, Italy

Francesco Mannella (francesco.mannella@istc.cnr.it)
Institute of Cognitive Sciences and Technologies, CNR
Via Romagnosi 18A, 00196, Rome, Italy

Luca Tummolini (luca.tummolini@istc.cnr.it)
Institute of Cognitive Sciences and Technologies, CNR
Via Romagnosi 18A, 00196, Rome, Italy

Abstract

Spatial demonstratives are deictic expressions used to point to a referent with language. In the standard view, they encode a spatial proximal\distal contrast between "near" (*this*) and "far" (*that*) from the speaker. Several studies have shown that such contrast maps on a perceptual contrast between peripersonal and extrapersonal space. Still, other factors beyond spatial distance influence demonstrative choice. Here we investigate whether the proximal/distal contrast maps also onto a more general contrast between being in physical control/not in control of a target referent. Participants were presented with two circles (red and blue) on a screen. They had to move them with the mouse to find the target circle (the one with two gaps). One circle followed the mouse trajectory (controllable), while the other moved randomly in the center of the screen (not controllable). Unknown to the participants, the gaps only appeared if the stimuli crossed a distance threshold. Importantly, participants had to use stimulus controllability to solve the task. They were instructed to answer by indicating the target to the experimenter using *this/that* and *red/blue* (in Italian *questo/quello* and *rosso/blu*). Results show that participants used the proximal demonstrative more frequently to refer to the target stimulus when in control. These findings suggest that, similarly to spatial distance, physical control influences demonstrative choice.

Keywords: language; semantics; physical controllability

Introduction

In the very first years of life, infants set the stage for the development of social cognition. In the emergence of cognitive mechanisms necessary for interacting with others and the world, the rise of joint attention is an important milestone. In this cognitively complex phenomenon, an actor, an addressee, and a referent are the three necessary elements. Indeed, we start talking about joint attentional focus at the end of the first year of life, when an improvement from a dyadic to a triadic interaction takes place (Eilan, 2005). At the same time, another triad is needed to create a joint attention situation. In this case, three cognitive elements: eye gaze, gestures, and language. Focusing on the last point, a specific class of grammatical markers is needed: demonstratives. For this reason, demonstratives are among the earliest words to be acquired (Clark, 1976; 2003). They are the most frequent words in adult lexicon (Leech & Rayson, 2014) and, even if with some differences across languages (binary contrast vs three term system), demonstratives are present in all languages (Diessel, 2006). Even if acquired so early, thus, they fulfill a complex role. According to Diessel (2006), demonstratives are spatial deictics with two functions:

1. They indicate the location of a referent in relation to a deictic center.
2. They coordinate the interlocutor's joint attentional focus.

In the standard view, when referring to a referent (e.g. an object with a certain location in space), people use the proximal demonstrative *this* (*questo* in Italian) when the object is located in the space near to the speaker. Conversely, the distal demonstrative *that* (*quello* in Italian) is preferred when the object is located far from the speaker.

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

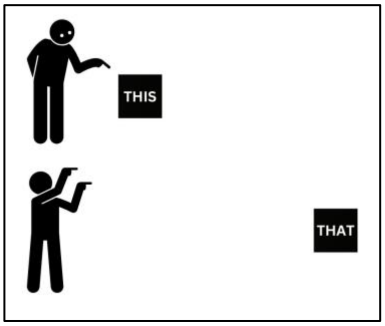

Figure 1: The standard view of the demonstrative choice.

The spatial hypothesis finds scientific support in Coventry and colleagues' studies, in which the choice *this/that* maps onto a distinction between peripersonal/extrapersonal space (Coventry, Valdés, Castillo & Guijarro-Fuentes, 2008; Coventry, Griffiths & Hamilton, 2014). Interestingly, just like tool use expands the peripersonal space boundaries (Farnè & Ladavas, 2000), tool use extends the region in space for which the proximal demonstrative is used (Coventry et al., 2008).

Still, spatial distance seems not to be the only factor driving demonstrative choice. The same authors provide evidence that other parameters to the choice: ownership, visibility, and familiarity (Coventry et al., 2014). Other authors, however, showed results in contrast with Coventry and colleagues' ones. For instance, Bonfiglioli and colleagues, showed that the proximal\distal contrast is present also within the boundary of the peripersonal space (Bonfiglioli, Finocchiaro, Gesierich, Rositani & Vescovi, 2009). In their study with Italian speakers, participants responded to instructions using demonstratives with a reach-to-grasp movement. Participants were faster in a congruent condition (*questo* used to refer to an object closer to the participant and *quello* for a more distant object), compared to the incongruent condition (*questo*-far, *quello*-close) even if the far object was always located in the peripersonal space. Interestingly, Cooperrider (2016) showed that demonstrative choice might be modulated by pointing gestures and their potential to reduce ambiguity. In particular, a more frequent use of *this* when pointing with the hand and in situations characterized by less ambiguity was shown. Interestingly, participants used *this* more frequently when pointing with the laser (less ambiguous) even far in space. Rocca and colleagues (Rocca, Tylén & Wallentin, 2019) showed that there is a preference for *this* to refer to objects more readily affording actions (smaller and harmless objects) regardless of any spatial parameter. However, even the basic assumption about demonstratives according to which demonstrative choice depends on egocentric proximity has been challenged by an EEG study (Peeters, Hagoort & Özyürek, 2019). Electrophysiological data, indeed, suggest that participants prefer the proximal demonstrative when an object is in a shared space between the speaker and an addressee, but only in the absence of another possible referent outside the shared space. In the presence of a potential reference outside the shared space participants don't show a preference for *this* or they even prefer the distal demonstrative.

In light of this evidence, the role of the peripersonal/ extrapersonal contrast in demonstrative choice is still unclear. Indeed, Kemmerer (1999) has argued that demonstrative choice maps in fact onto a more abstract space rather than a concrete and physical space. For instance, consider a sentence like "this planet is smaller than that planet" uttered when looking at the sky: clearly the proximal\distal contrast felicitously applies well into the extrapersonal space. But then, which of these views can better explain demonstrative choice? Or alternatively can they be made consistent with each other?

With the aim of providing a unifying perspective, Brovold and Grush (2012) proposed a theory that might in fact represent a bridge between these alternatives. In their proposal, demonstrative choice leverages on the speakers' awareness of their own controllability of the referent. In this view, demonstrative semantics (both production and understanding of spatial demonstratives) relies on the interlocutors' spontaneous assessment of their own and each other's "control domains": which objects they have control over and which they don't. In other words, the proximal/distal contrast instead of reflecting a more superficial spatial one maps onto a deeper (and abstract) contrast between in control/not in control. While Brovold and Grush distinguish three types of control, here we focus on physical control and leave other kinds for future investigation.

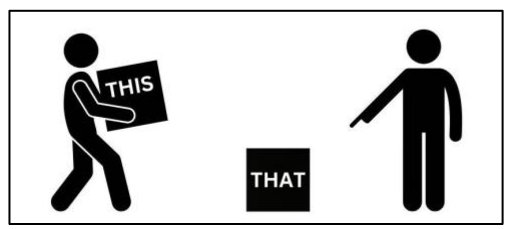

Figure 2: Demonstrative choice based on physical controllability.

While this theoretical framework has the potential to accommodate much evidence on spatial demonstratives, direct experimental support has not been so far provided. Thus, the aim of this study is to experimentally test the hypothesis according to which, when people are in physical control of an object, they tend to refer to it using the proximal demonstrative (*this*) and, vice versa, when they are not in control of the object they prefer the distal one (*that*).

Method

To assess the effect of physical control on demonstrative choice, we adapt the visual search task of Wen and Haggard (2018) in which participants use the mouse to interact with stimuli with different level of physical controllability. We elicit demonstratives without participants being aware that their linguistic choice is relevant for the purposes of the experiment. The experiment was run in Italian.

Participants

In line with previous experiments carried out by Coventry, Griffiths & Hamilton (2014), we recruited twenty participants for this study. All participants were healthy, right-handed and native Italian speakers (11 women, mean age = 25,76, SD = 4,21) and not bilingual from birth. All participants used their right hands to move the mouse during the tasks. All participants had corrected-to-normal visual acuity. The exclusion criterion was 75% of the valid trials (i.e., trials with correct target identification). No participant met this criterion. Participants gave their written informed consent before the experiment started.

Procedure

Participants were asked to sit at a desk in front of a screen. To avoid differences in terms of distance from the screen, we ensured that, for each participant, the screen was located at the boundary between each participant's peripersonal and extrapersonal space. To do that, we asked participants to extend the arm in front of them, towards the screen, and we moved their chair location so that, with their arm fully extended, the tips of their middle finger was 20cm far from the screen. Importantly, this procedure also ensures that all the stimuli were always located in the extrapersonal space of participants. The experimenter sat next to the participant to simulate a condition of joint attention with minimal interference with the task. The experimenter randomly changed sides across participants to avoid any possible side bias. Once in position, we asked participants to fill the informed consent. Handedness was assessed with the Edinburgh handedness inventory (Oldfield, 1971). As in Coventry et al. (2008), participants were not informed about the real aim of the experiment and any reference to the relevance of spatial demonstratives for the experimental study was avoided. Participants were given a visual search task and were told to communicate their choice to the experimenter who sat either on their left or on their right. To limit linguistic variability across participants, they were instructed to linguistically point to the relevant target using only the Italian demonstrative system: "questo" or "quello". Participants read the instructions on the screen and if they had any doubts, they could ask the experimenter for further instructions. After the instructions, participants had the possibility to familiarize with the task in a practice session.

In the visual search task, participants had to identify, as fast as possible, the target stimulus and to report their selection to the experimenter. The entire experimental session was composed of two sessions interspersed with a break. Each session was composed of six blocks and each block was composed of eight trials for a total of 96 trials. All trials began with a fixation cross (0.5s) on a white screen followed by two stimuli (see Figure 3). At the start of each trial, two circles (with a diameter of 1 cm), with a colored circumference (one red and one blue) appeared on the screen. The center of the circles was not filled (white). Both circles stayed on the screen for all the duration of the trial (5s). Once participants started to move the mouse, the two circles started to move as well but the relation between the mouse movements and the trajectory of the motion of the circles could vary. More specifically, in each trial, one of the two stimuli had a high level of physical controllability (90% control), while the other one had a low level of physical controllability (10% control). For the high controllable stimulus, the other 10% (and for the low controllable the other 90%) of the trajectory was randomly drawn from a database of trajectories acquired by the authors before the beginning of the data collection. In particular, while the high controllable circle could be moved in each part of the screen by moving the mouse, the low controllable stimulus mostly stayed in a given area (around the center of the screen). Only when participants started to move the mouse, the shape of the circles changed. One of the two circles was the distractor (one gap in the circumference) while the other one was the target stimulus (two gaps in the circumference). Moreover, in each trial the color of the circle was randomized so that in half of the trials the target stimulus was red and in the other half it was blue. Importantly, unknown to the participants, the shape of the circles changed only when the two stimuli reached a certain distance in the screen. In other words, to solve the task more efficiently (to detect the target stimulus), participants could discover the underlying strategy and exploit the controllability of one of the two circles in their favor. By moving the high controllable circle far from the other (the low controllable circle), participants could more efficiently discover the perceptual feature. Next, participants were required to respond in Italian by saying either "questo" or "quello" to inform the experimenter of their choice. After the five seconds of the stimuli presentation a new screen appeared on the monitor showing a red and a blue dot. In this case, participants had to indicate what was the color of the target stimulus (in Italian "rosso" or "blu"). As in Coventry and colleagues' study (Coventry et al., 2008), in each trial the experimenter noted the linguistic choices made by participants. To increase salience of the relative controllability of the stimuli, an attentional check at the end of each block of eight trials was added. In the attentional check, we asked participants "in the previous trial, did you control the target stimulus?". They had to answer to this question by pressing the "S" to say yes and "N" to say no.

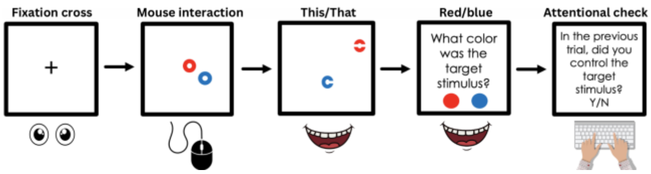

Figure 3: Timeline of a trial in the experiment.

Results

We computed the percentage of demonstrative choices of proximal *questo* or distal *quello* (see Figure 4). We ran a generalized linear mixed model on the proportion of proximal choices with "Target" (target/not target) as fixed factor and "Subject" as random factor. From this model, a significant effect of controllability of the target stimulus on the use of "this" has emerged (Z= 17.187, p<.001).

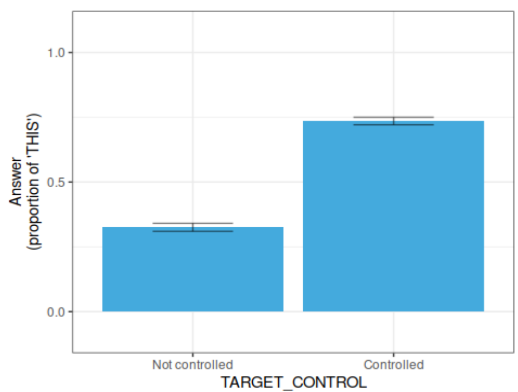

Figure 4: Proportion of proximal choices used to refer to the target stimuli.

The proportion of proximal choices was significantly higher when the target stimulus was controlled compared to when it was not controlled.

Next, we run a second generalized linear mixed model in order to assess whether proximal choice changed as a function of the block: the interaction between Target (target/not target) and Blocks (the number of the block). In this case as well, a main effect effect of controllability on proximal choice was also observed (Z=6.399 p<.001). We also observed a significant interaction with blocks (Z= 3.222, p=.001) suggesting in particular that the propensity to use the proximal demonstrative when the target was not controlled decreases in time.

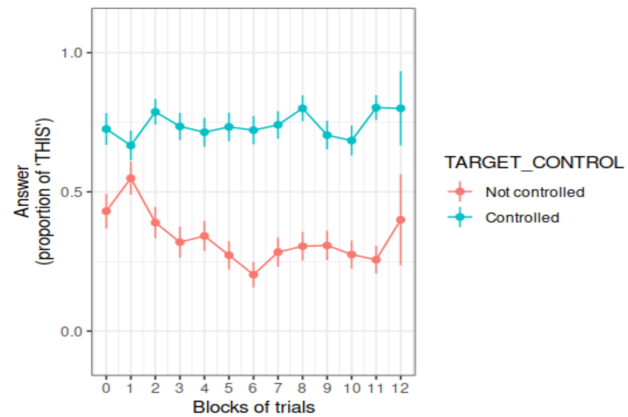

Figure 5: Proportion of proximal choices to refer to the target stimulus along blocks.

Finally, we conducted an exploratory analysis on movement trajectories. We reasoned that different motor strategies can reveal whether participants exploit stimulus controllability to solve the visual search task or adopt a more exploratory strategy of random movements. We assumed that goal-directedness implies a stronger sense of control, and conjectured the adoption of a goal-directed strategy could influence demostrative choice. Participants engaged in 96 trials per session, with each trial lasting 296 seconds. During these trials, positional data was continuously recorded at one-second intervals, generating a time series of 296 data points for each trial. We examined the data as one-dimensional time series, focusing solely on the distance from the center of the screen as our primary reference point because it was the only direction of movement that led to change the shape of the circles. We conducted an analysis of these trajectory time-series using principal component analysis (PCA). The initial two components effectively elucidated 67% of the total variance. The plot in Figure 6 visually represents the time series projected trajectories across the initial 8 components

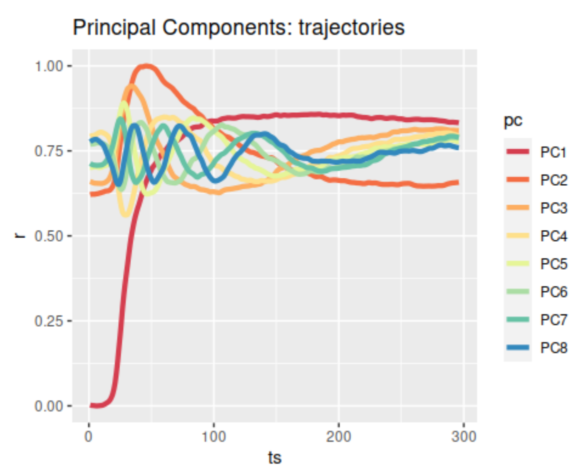

Figure 6: Principal components of the mouse trajectories.

The first two components capture two distinct behavioral patterns. The first component involves displacing the controllable stimulus from the center of the screen towards the borders. The second component involves an initial movement towards the borders in the first phase of the trial, followed by a return to the center in the subsequent phase. Each trial within every session was then categorized by identifying the component with the strongest correlation to the trial time series. A trial received the label 'PC1' if the first principal component had the strongest correlation, 'PC2' if the second principal component had the strongest correlation, and so on. We correlated the probability of each component across subjects (the proportion of trials labeled with that component), with the participant's response sensitivity to target. Sensitivity to target control was calculated by assigning a sensitivity score of 1 if the subject always responded *questo* when the target was in control or *quello* when it wasn't. Conversely, if the subject responded *quello* when the target was in control or *questo* when it wasn't, the sensitivity score was set to 0.

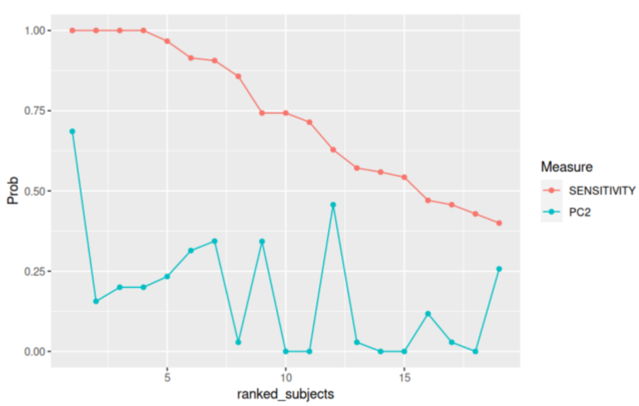

Figure 7: Classification of participants according to sensitivity to target control (in red) and probability to adopt PC2 (in blue).

On this basis, we explored the correlation between sensitivity and the probability of each component across subjects. We did not find significant correlations between the different principal components and target control sensitivity. However, the only correlation that was close to significance was with the PC2 component that describes a strategy of initially moving the object towards the border of the area and then moving it back to the center of the screen that is suggestive of a goal-directed strategy (Pearson's product-moment correlation, r=0.45, t(17)=2.0705, p=.05395). In general behavioral strategies are not correlated with participants' sensitivity to control.

Discussion

Our study provides novel evidence in the debate on the semantics of spatial demonstratives providing empirical support for an effect of physical control over a referent on the speaker's demonstrative choice. Results of this study are therefore consistent with Brovold and Grush's theory (Brovold & Grush, 2012), who hypothesize that selection of which demonstrative to use to refer to a target takes into account the speaker's physical control over the referents (or his\her lack thereof). In our study, we leverage a visual search task in which participants had to move two stimuli (one controllable and one non-controllable and, at the same time, a target vs a non-target circle) on the screen with their mouse. To solve the task, participants could exploit the physical controllability of one of the two circles to make target perceptual feature appear more efficiently. We show that, when Italian participants\speakers intended to indicate to the experimenter which stimulus was the target one, they preferred to use the proximal demonstrative *questo* ("this") when in physical control of the target stimulus. Consistently, they reveal a preference for the distal demonstrative *quello* ("that") when lacking control over the target stimulus. Moreover, our results show that this tendency increases over the blocks. We interpret this result in part as a learning effect, and in part in terms of increasing awareness on stimulus controllability over time. The latter aspect suggests that reducing ambiguity is indeed an important factor as proposed by Cooperrider (2016). Instead of simply adding evidence for a new non-spatial parameter, in line with Brovold & Grush (2012), we suggest that these results might provide support for a more general explanation. Spatial parameters are strictly interconnected with physical controllability considering that usually the near space entails physically controllable and that physical controllability characterize objects that afford interaction (Rocca, Tylén & Wallentin, 2019). Still physical control can also be extended far in space too as when remotely control a flying drone.

The importance of physical controllability in the demonstrative choice shares new lights on the link between these universal linguistic systems and action. We want to stress, thus, that the demonstrative choice is not just a matter of space perception, but that the possibility to act towards a referent and to do it in joint attentional situations plays a key role. Our view, thus, is in line with the one according to which language may have evolved from actions and gestures in human evolution (Arbib, 2012) as well as in children's development (Iverson & Goldin-Meadow, 2005). Considering, indeed, that demonstratives are the earliest words that we acquire, our results might suggest that their function is strictly interconnected with the possibility, in children, to be able to physically control and, consequently, interact with referents in space and the world that, otherwise, would not be physically accessible.

While results of this study supports this view, it is important to acknowledge several limitations, which can be addressed in future work. Firstly, we have only focused on the exophoric use of demonstratives in the extra-linguistic context leaving aside their intralinguistic use (endophoric reference). Secondly, this study has collected evidence for

Italian only and generalization to other languages may not be ensured. Indeed, although the demonstratives contrast relies on a binary distinction in Italian as well as in English, in other languages it relies on a three terms system as in Spanish (Diessel, 2005). Thirdly, the analysis of movement trajectories did not reveal an association between behavioral strategies and their sense of control (operationalized as sensitivity to target control). One possible reason for this lack of evidence could be due to the experimental design. In the current paradigm, the explicit task of discriminating between the shapes of the circles when they changed hid an instrumental task, which is to find a way to change the shape of the stimuli by using the stimulus under control. However, in the current version the instrumental task can be easily solved with many different behavioral strategies. As a result, it may not be straightforward to determine from behaviors whether participants exploited the controllability of the stimulus to solve the task or simply moved randomly. A future direction of this study involves refining the paradigm so that the underlying instrumental task can be used as a covariate measure of the participants' sense of physical control. Finally, while Brovold & Grush (2012) proposed a broader theory about controllability, taking in consideration other forms of control beyond the physical (like perceptual and social control), in this study, we just focused on physical control. For these reason, future work will replicate and extend these findings with new paradigms to investigate perceptual and social control. These possibilities have the potential to provide new evidence about other parameters already highlighted in previous studies such as ownership, visibility, and familiarity.

Acknowledgments

We thank the anonymous reviewers for useful comments and acknowledge funding by PRIN2022 "Understanding What is Mine" grant No 2022LYRT8E and PRIN-PNRR2022 "Normativity of Joint Action" grant no. P2022YYRK3.

References

Arbib, M.A., (2012). *How the brain got language: The mirror system hypothesis* (Vol. 16). Oxford University Press.

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.

Brovold, A., Grush, R. (2012). Towards an (improved) interdisciplinary investigation of demonstrative reference. *Perception, realism, and the problem of reference*, 11-42.

Clark, E. V. (1976). From gesture to word: On the natural history of deixis in language acquisition. *Human growth and development*, 1-38.

Clark, E. V. (2003). Acquiring language: Issues and questions. *First language acquisition*, 1-21.

Cooperrider, K. (2016). The co-organization of demonstratives and pointing gestures. *Discourse Processes*, 53(8), 632-656.

Coventry, K., 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., Griffiths, D., & Hamilton, C.(2014). Spatial demonstratives and perceptual space: Describing and remembering object location. *Cognitive psychology*, 69,46.

Diessel, H. (2005). Distance contrasts in demonstratives. World atlas of language structures, 170-173.

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

Eilan, N. (Ed.). (2005). *Joint attention: Communication and other minds: Issues in philosophy and psychology*. Oxford University Press, USA.

Farnè, A., & Làdavas, E. (2000). Dynamic size-change of hand peripersonal space following tool use. *Neuroreport*, 11(8), 1645-1649.

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological science*, 16(5), 367-371.

Kemmerer, D. (1999). "Near" and "far" in language and perception. *Cognition*, 73(1), 35-63.

Leech, G., & Rayson, P. (2014). *Word frequencies in written and spoken English*. Routledge.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113.

Peeters, D., Hagoort, P., & Özyürek, A. (2015). Electrophysiological evidence for the role of shared space in online comprehension of spatial demonstratives. *Cognition*, 136, 64-84.

Rocca, R., Tylén, K., & Wallentin, M. (2019). This shoe, that tiger: Semantic properties reflecting manual affordances of the referent modulate demonstrative use. *PloS one*, 14(1).

Wen, W., & Haggard, P. (2018). Control changes the way we look at the world. *Journal of cognitive neuroscience*, *30*(4), 603-619.