



HAL
open science

Human-Human Guidance Study

Kathleen Belhassein, Aurélie Clodic, Hélène Cochet, Marketta Niemelä, Päivi Heikkilä, Hanna Lammi, Antti Tammela

► **To cite this version:**

Kathleen Belhassein, Aurélie Clodic, Hélène Cochet, Marketta Niemelä, Päivi Heikkilä, et al.. Human-Human Guidance Study. Rapport LAAS n° 17596. 2017. hal-01719730

HAL Id: hal-01719730

<https://laas.hal.science/hal-01719730v1>

Submitted on 28 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

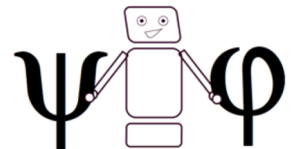
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HUMAN-HUMAN GUIDANCE STUDY

Report v1.0



MuMMER



JointAction4HRI

Kathleen BELHASSEIN^{1,2}

Aur lie CLODIC¹

H l ne COCHET²

Marketta NIEMEL ³

P ivi HEIKKILA³

Hanna LAMMI³

Antti TAMMELA³

¹LAAS-CNRS, Toulouse, France

²CLLE, Toulouse, France

³VTT, Espoo, Finland

INTRODUCTION

The present study aims to analyze the elements of a guiding situation in the Ideapark shopping center, the final project being to integrate the robot Pepper by Softbank Robotics as a user guide. To make the customers feel comfortable with the guidance service and understand the guidance of the robot, we first need to understand the keys of an effective human-human guiding situation, by observing the experimented guides indicate several paths to take to customers.

Context

To describe a route to someone, we separated it into a set of segments that connect important landmarks and a set of actions (Kopp et al., 2007). We speak of vectors for the entities where the actions must be carried out, like the roads (Michon & Denis, 2001). The guide can also use particular points within these vectors, used to indicate when exactly an action should be performed ("at the crossroads"). The guide can choose to use a route perspective (e.g. indicate the path by taking the point of view of the subject he is guiding) or a survey perspective (i.e. birds-eye view) (Striegnitz et al., 2009). Finally, to help the subjects to find their bearings, the guide may use landmarks (Denis, 1997), i.e. objects along the vectors.

In a guiding situation, humans use more gestures than in a classic conversation. The gestures used to describe a route to take can be classified according to McNeill's typology (1992). In this situation, we will focus mainly on deictic gestures, iconic gestures, metaphorical gestures, and finally beats. Deictic gestures refer to pointing gestures, i.e. a way to direct the attention of our interlocutor towards a specific object. The iconic gestures are gestures that visually illustrate the morphological characteristics of the object described. Metaphorical gestures are gestures that abstractly represent a concept (e.g. "Stay calm", accompanied by a gesture of a slow beat of the hand horizontally). Finally, the beats are gestures which punctuate speech but without any semantic content (Tellier, 2014). In the case of a guiding situation, deictic gestures will naturally be more frequent than iconic gestures or beats, while metaphorical gestures will be rare (Allen, 2003).

Children tend to use the ipsilateral hand in the referent's visual field (Cochet, Jover, Oger & Vauclair, 2014). However, many studies have shown a bias favoring the use of the right hand in communicative gestures in adults (Kimura, 1973; Bryden, Pryde & Roy, 2000), in children (Bates et al., 1986; Hannan & Fogel, 1987; Cochet & Vauclair, 2010), and even in primates (Hopkins et al., 2005; Meguerditchian & Vauclair, 2009). Thus, we would be more likely to

point with the right hand in the left visual field, than with the left hand in the right visual field (Butterworth et al., 2002; Esseily, Jacquet & Fagard, 2011). Moreover, it is important to emphasize that this is not necessarily correlated with the hand used preferentially for manipulative actions (Hopkins et al., 2005; Vauclair & Imbault, 2009; Esseily, Jacquet & Fagard, 2011).

The pointing gestures can take two different forms: index, or open palm. It would seem that the open palm is rather used when there is a little ambiguity to identify the referent while we will preferentially point with the index when we perceive an ambiguity (Wilkins, 2003; quoted by Cochet & Vauclair, 2014). We therefore point with the open palm big referents as locations because they require less precision and are more easily discriminated. On the other hand, smaller targets or targets in presence of several similar objects in the proximal environment make them more difficult to discriminate and will therefore be pointed with the index (Cochet, Jover, Oger & Vauclair, 2014). In the same way, the distance with the referent influences the shape of the pointing gesture. Objects that are not visible or far away are more likely to be pointed with the palm open, while close referents will be associated with an index gesture (Cochet & Vauclair, 2014).

Other studies estimate that the index will be used to indicate landmarks or objects, most often visible by the communicative partners, or when the pointed object is the main objective of the speech (Mechraoui & Noor Mohd Noor, 2015). The open palm, meanwhile, will be preferred to indicate the paths or turns. It is also possible to use both hands: in this case, the gestures will be in many cases identical (Cochet & Vauclair, 2014).

The orientation of the hand also varies according to a lot of factors, and studies have shown a link between the morphological characteristics of the referents described and those of the gestures used. For example, in a study on route description, 85% of gestures performed with the open palm and laterally oriented referred to a directed action (Kopp et al, 2007). It has also been shown that an orientation of the hand facing the ground will be preferred for horizontal referents, while a lateral orientation of the hand will be used for vertical referents as stores (Cochet & Vauclair, 2014). Finally, it is important to emphasize that the orientation of the palm and the shape of the gesture are dependent on the context and the subject (Cochet & Vauclair, 2014).

The mutual gaze is a crucial element of any cooperation situation; it allows us to verify that our interlocutor listens, understands, and is attentive to what we communicate to them (Kendon, 1967). Thus, it can significantly improve collaborative performance (Neider et al., 2010). In a guiding situation, the mutual gaze occurs at several moments: at the beginning and

at the end of the interaction, during the initial question, or at the end of the explanations of the guide (Boot, 2005). Most of the time, the gaze, the position, the gestures and the description all start at the same time for the first mentioned location. If the guide is sure of his description, he keeps the mutual gaze; on the other hand, we observe much less mutual gaze when this one doubts. Finally, the gaze naturally follows the arm's gesture.

Studies have shown the importance of the gaze in human-robot interaction, and in particular the fact that it significantly improves human performance (Staudte & Crocker, 2009). In a human-robot study, participants were asked to identify what the robot (PR2) was pointing at. The results show that participants used the direction of the arm to interpret the pointing of the robot, but also and especially the orientation of the head (gaze) when it was in the same direction as the arm (Williams et al., 2013). Similarly, a study shows that humans can anticipate the location of a target before it is specified in the speech, thanks to the direction of the gaze of the iCub robot (Boucher et al., 2012).

THE STUDY

The human-human guidance study was conducted at Ideapark shopping center in September 2017 by VTT. Two persons working at the information desk of Ideapark were asked to take part in video recordings, where researchers asked guidance to different places in the shopping center. The questions were made beforehand by researchers, based on the earlier interview of another person working at the information desk (conducted in the autumn 2016 by VTT). Moreover, various criteria were taken into account to choose the locations that will be requested by the customer: it was indeed important that all directions of Ideapark are covered with locations near or farther from where the guide and the customer are located. The locations could be in front or behind the guide or the customer, and being visible or not. A question could cover several possible locations (e.g. "shoe shops"). One of the situations was also that the customer could request the location of two different places in the same question (e.g. "I'm going to the pharmacy and then to Minimani"). Finally, two questions in English were also asked.

The study was conducted at the assumed future location of Pepper (which has been changed a bit after the study), where a square of 4m² (2m x 2m) was marked in the floor. The guide was asked to stay inside the square, and guide the researchers to a desired place in their natural way, with speech and gestures (as they would normally do at the information desk), without using a map. The researchers approached the guide from different directions, which made the guide to change her orientation. The guiding situations were audio recorded and

videotaped with a top camera. A third researcher asked passers-by to stay outside the videotaped area (according to the policy of VTT, which does not allow customers to be videotaped without a permission asked from them). After video recordings, the guides were briefly interviewed. The interviews focused especially on the landmarks used to facilitate guidance.

The both guides were used to guide customers in the shopping center. The first of the guides (in the video recordings) had worked at the information desk of Ideapark for several years, and the second guide for several months.

At VTT, the video recordings were edited to be transferred to LAAS, the audio recordings were transcribed and translated from Finnish to English, and the interviews with the guides were reported. In addition, examples of landmarks used in the guidance were photographed and added to the report of the study.

OBSERVATIONS

Gestures

In these different guiding situations and in agreement with Allen (2003), we can notice that deictic gestures are much more frequent than iconic gestures or beats, while there are less metaphorical gestures. In the various interactions of the two guides, we observe the preferential use of the ipsilateral hand to the visual field of the referent, but also and mostly the preferential use of the right hand. On the other hand, we can sometimes see the use of the left hand when the guide indicates a left turn and this, even if he used the right hand until then.

There is also a relatively frequent use of the open palm. This seems invariant with Guide1, who almost not use the index to point, unlike Guide2. Based on the articles cited above, we can assume that this is because the referents are locations, stores, which therefore requires less precision in their size. We can nevertheless notice the occasional use of the index in the case of referents not visible by the guide and the subject (image 1), more particularly for Guide2. This is also the case for most of the iconic gestures, used to describe landmarks on the way to the requested location.

With respect to the orientation of the palm, the observations are consistent with the data from the literature. We thus observe a preferential use of the lateral position (image 1) for the two guides (and more particularly with the open palm) to indicate the stores (vertical referents) or directed actions like the paths to be taken and the turns. The position and shape of Guide1's gestures are almost invariant (palm oriented laterally). On the other hand, when the guide refers

to a horizontal referent and especially here to the escalator, the orientation of the hand will also be horizontal (palm facing the ground).



Image 1: Illustration of some different forms and orientations of the gestures. On the left, the open palm is oriented laterally. At the center, the guide point with the index, also laterally oriented. On the right, the open palm is oriented horizontally.

When the guides use both hands simultaneously, the actions performed are actually identical but can be performed not exactly in synchrony. We also note the use of both hands for some metaphorical gestures (image 2).

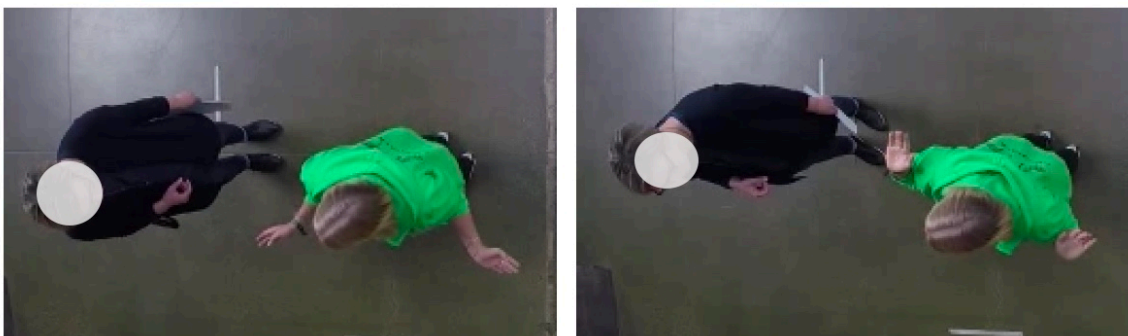


Image 2: The guide uses her both hands simultaneously. On the right, the guide uses a metaphorical gesture with her both hands.

When the guide wants to insist on the referent's distance (e.g. when the referent is located at the end of a corridor, or if it is the last store ...), the signal is adjusted and the guide presses on his gesture, "waving" his hand up and down.

Finally, once the question is asked by the subject, we observe that the guide tends to point first to the direction of the referent itself (survey perspective), then indicate the path to access it (route perspective), sometimes with landmarks when the referent is far and difficult to access.

Positions

The fact that the guide preferentially uses the right hand gives rise to a position which favor the use of this hand: when the guide has to move to point a referent (e.g. the requested location or a landmark behind his back), he will most often go in profile to have the referent on his right and the subject on his left. So, he can point with his right hand. This observation is

very frequent with Guide1, and almost systematic. When the path indicated by the guide is complex (referent not visible, far, with obstacles ...), the guide and the subject can sometimes get closer to allow the subject to see from the point of view of the guide and therefore precisely see the referent pointed (image 3).



Image 3: The guide and the subject get closer

Gazes

Regarding the alternation of gaze, we can only rely on the gaze of the guide. Indeed, the previously defined scenario followed by the subject and the fact that he does not really want to go to the requested location thus prevent a natural interaction. Moreover, given the impossibility of recording the faces of the subjects, we can only infer the gaze from the orientation of their head.

We observe that a mutual gaze takes place during the question of the subject. The guide checks the subject's understanding by directing his gaze towards him at the end of each of his descriptions, or part of his descriptions. In agreement with the observations of Lenart Boot (2005), the gaze, the position, the gestures and the description of the guide often all start at the same moment during the first description of the location. We also observe with Guide2 a rupture of the mutual gaze when this one doubts on the location of the referent asked: Guide2 does not look any more at the subject in this situation.

On the other hand, we observe more alternation of gaze when the guide is placed in profile with respect to the referent and the subject: his gaze then goes back and forth between them. The mutual gaze also appears at each metaphoric gesture. Iconic gestures and beats are also often (but not always) along with mutual gaze. Finally, we note that the more the referent is far and difficult to access, the more there are mutual gazes; especially because the description of the path is longer in these cases.

Other observations

We can notice that Guide2 uses much more landmarks than Guide1, whereas Guide1 use metaphoric gestures when Guide2 do not. In reality, we find similarities between the two guides (as we see for example on Image 2), but we also observe subject-dependent differences at the level of the gestures used (shape, orientation, type), the mutual gaze and the use of landmarks. Surprisingly, the most used landmarks are not the big stores (e.g. Prisma), but the doors (e.g. A door), “Central Park” (a big square at the center of the mall) and the “old town” (section on Ideapark just behind the info desk). Furthermore, the guides don’t use street names to indicate the path to take, because they are not clearly visible. The guides use landmarks in their speech when the referent is not visible from where they are, and when they explain to the subject the path to get to the desired location (route perspective). Landmarks are indeed much more rarely used when they indicate the direction of the referent directly (survey perspective).

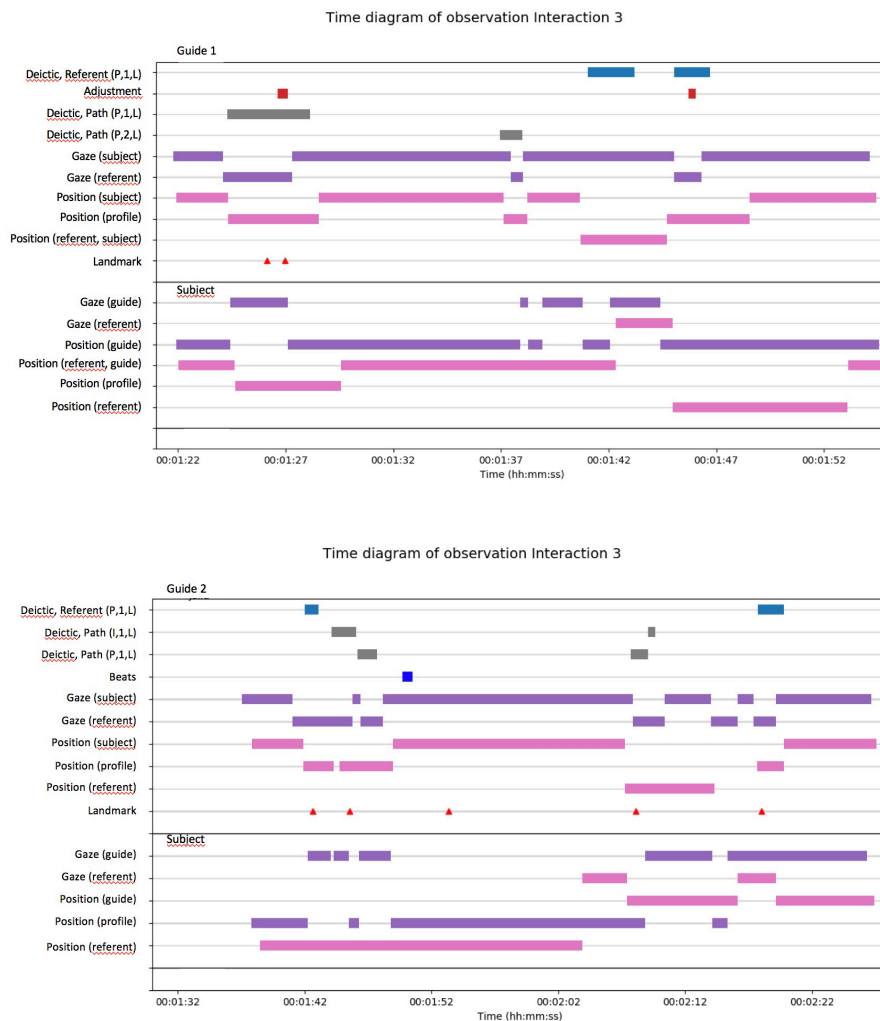


Image 4 : Time diagrams of the same interaction for Guide1 (up) and Guide2 (down). We can observe some similarities in gazes and positions. We also notice the use of landmarks.

CONCLUSIONS AND PERSPECTIVES

This study aimed to understand the key elements of a guiding situation. We observed gestures, positions, gazes or landmarks used by the two guides. Although non-verbal communication in a guidance situation is subject-dependent and context-dependent, we have nevertheless been able to find patterns of behavior that are consistent with literature on gesture studies.

The main observation in this study was the use of the ipsilateral hand in the visual field of the referent as well as the preferential use of the right hand, which is consistent with the data of the literature. We thus observed that the guide chooses a position in the space which will favor the use of the right hand (the guide will be positioned to have the referent to point to his right and the subject to his left). We also noticed that many of the gestures were performed with the open palm laterally oriented, to refer to either a vertical target (such as a store) or a directed action (typically when the guide shows the path to get to the requested location). Finally, we observed that the mutual gaze appeared at different moments of the interaction, especially at the end of each part of the guide's description of the path. The mutual gaze is all the more frequent as the requested location is far away, complicated to access, and as the guide is placed in profile regarding the subject and the target.

However, the main limit of these videos is that the subject was reading a scenario, which leads to a biased interaction. Since the subject does not really want to go to the requested location and reads a text that asks him to position himself in a certain way in relation to the guide, and that tells him all the sentences to say, he does not interact naturally with the guide. The mutual gaze is particularly affected by this bias, because the lack of concentration of the subject on the explanations of the guide.

We would therefore need a situation where the subject really has an interest in going to the requested location, and without proposing him a prepared sentence. With a more natural situation, we will be able to see the gazes' alternations between the subject and the guide, and most importantly what is happen when the subject does not understand the explanations of the guide. In a future study, we will therefore be interested in the modifications of the guide's gestures in this particular situation.

Acknowledgements This work was supported by the European Union's Horizon 2020 research and innovation programme under grant agreement no. 688147 (MuMMER) and a

French National Research Agency (ANR) grant, reference ANR-16-CE33-0017 (JointAction4HRI).

REFERENCES

- Allen, G. L. (2003). Gestures Accompanying Verbal Route Directions: Do They Point to a New Avenue for Examining Spatial Representations? *Spatial Cognition & Computation*, 3(4), 259-268. https://doi.org/10.1207/s15427633scc0304_1
- Bates, E., O'Connell, B., Vaid, J., Sledge, P., & Oakes, L. (1986). Language and hand preference in early development. *Developmental Neuropsychology*, 2(1), 1-15.
- Boot, L. (2005). Gaze during the giving of directions (Student Work). University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science.
- Boucher, J.-D., Pattacini, U., Lelong, A., Bailly, G., Elisei, F., Fagel, S., ... Ventre-Dominey, J. (2012). I Reach Faster When I See You Look: Gaze Effects in Human–Human and Human–Robot Face-to-Face Cooperation. *Frontiers in Neurobotics*, 6. <https://doi.org/10.3389/fnbot.2012.00003>
- Bryden, P. J., Pryde, K. M., & Roy, E. A. (2000). A Performance Measure of the Degree of Hand Preference. *Brain and Cognition*, 44(3), 402-414. <https://doi.org/10.1006/brcg.1999.1201>
- Butterworth, G., Franco, F., McKenzie, B., Graupner, L., & Todd, B. (2002). Dynamic aspects of visual event perception and the production of pointing by human infants. *British Journal of Developmental Psychology*, 20, 1-24.
- Cochet, H., Jover, M., Oger, L., & Vauclair, J. (2014). Morphological Differences Between Imperative and Declarative Pointing: Hand Shape, Arm Extension, and Body Posture. *Journal of Motor Behavior*, 46(4), 223-232.
- Cochet, H., & Vauclair, J. (2010). Features of spontaneous pointing gestures in toddlers. *Gesture*, 10(1), 86-107. <https://doi.org/10.1075/gest.10.1.05coc>
- Cochet, H., & Vauclair, J. (2014). Deictic gestures and symbolic gestures produced by adults in an experimental context: Hand shapes and hand preferences. *Laterality: Asymmetries of Body, Brain and Cognition*, 19(3), 278-301. <https://doi.org/10.1080/1357650X.2013.804079>
- Denis, M. (1997). The Description of Routes: A Cognitive Approach to the Production of Spatial Discourse. *Current Psychology of Cognition*, 16, 409-458.

- Esseily, R., Jacquet, A.-Y., & Fagard, J. (2011). Handedness for grasping objects and pointing and the development of language in 14-month-old infants. *Laterality: Asymmetries of Body, Brain and Cognition*, 16(5), 565-585. <https://doi.org/10.1080/1357650X.2010.499911>
- Hannan, T. E., & Fogel, A. (1987). A Case-Study Assessment of 'Pointing' during the First Three Months of Life. *Perceptual and Motor Skills*, 65(1), 187-194. <https://doi.org/10.2466/pms.1987.65.1.187>
- Hopkins, W. D., Russell, J. L., Freeman, H., Buehler, N., Reynolds, E., & Schapiro, S. J. (2005). The distribution and development of handedness for manual gestures in captive chimpanzees (*Pan troglodytes*). *Psychological Science*, 6, 487-493.
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica (Amst.)*, 26, 22-63.
- Kimura, D. (1973). Manual activity during speaking— I. Right-handers. *Neuropsychologia*, 11(1), 45-50. [https://doi.org/10.1016/0028-3932\(73\)90063-8](https://doi.org/10.1016/0028-3932(73)90063-8)
- Kopp, S., Tepper, P. A., Ferriman, K., Striegnitz, K., & Cassell, J. (2007). Trading Spaces: How Humans and Humanoids Use Speech and Gesture to Give Directions. In T. Nishida (Éd.), *Wiley Series in Agent Technology* (p. 133-160). Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470512470.ch8>
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: The University of Chicago Press.
- Mechraoui, A., & Noor Mohd Noor, F. (2015). Pointing Gestures in Direction Giving Interactions of Malaysian English Speakers. Presented at Proceedings of the APAN - Network Research Workshop 2015.
- Meguerditchian, A., & Vauclair, J. (2009). Contrast of hand preferences between communicative gestures and non-communicative actions in baboons: Implications for the origins of hemispheric specialization for language. *Brain and Language*, 108(3), 167-174. <https://doi.org/10.1016/j.bandl.2008.10.004>
- Michon, P.-E., & Denis, M. (2001). When and Why Are Visual Landmarks Used in Giving Directions? (p. 292-305). Presented at International Conference on Spatial Information Theory, COSIT 2001.
- Neider, M. B., Chen, X., Dickinson, C. A., Brennan, S. E., & Zelinsky, G. (2010). Coordinating spatial referencing using shared gaze. *Psychonomic Bulletin & Review*, 17, 718-724.
- Staudte, M., & Crocker, M. W. (2009). Visual attention in spoken human-robot interaction (p. 77). ACM Press. <https://doi.org/10.1145/1514095.1514111>

- Striegnitz, K., Tepper, P., Lovett, A., & Cassell, J. (2009). Knowledge Representation for Generating Locating Gestures in Route Directions. In K. R. Coventry, T. Tenbrink, & J. Bateman (Éd.), *Spatial Language and Dialogue* (p. 147-165). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199554201.003.0011>
- Tellier, M. (2014). Quelques orientations méthodologiques pour étudier la gestuelle dans des corpus spontanés et semi-contrôlés. *Discours*, (15). <https://doi.org/10.4000/discours.8917>
- Vauclair, J., & Imbault, J. (2009). Relationship between manual preferences for object manipulation and pointing gestures in infants and toddlers: Handedness for object manipulation and pointing. *Developmental Science*, 12(6), 1060-1069. <https://doi.org/10.1111/j.1467-7687.2009.00850.x>
- Wilkins, D. (2003). Why pointing with the index finger is not a universal (in sociocultural and semiotic terms). In S. Kita (Éd.), *Pointing: Where language, culture, and cognition meet* (p. 171-215). Mahwah, NJ: Erlbaum.
- Williams, M.-A., Abidi, S., Gärdenfors, P., Wang, X., Kuipers, B., & Johnston, B. (2013). Interpreting Robot Pointing Behavior. In G. Herrmann, M. J. Pearson, A. Lenz, P. Bremner, A. Spiers, & U. Leonards (Éd.), *Social Robotics* (Vol. 8239, p. 148-159). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-02675-6_15