ADEALO: Au Doigt Et A L'Œil

Metrics for the grammar of deictic pointing

Adrian VOICU, Master "AI 4 One Health"
Under the supervision of Christian Graff[1] and Ernesto Gomez[2]
With assistance from Jade Chomat, Laurine Reynas, Louison Barbe and Coline Fons

[1] LPNC (Body and Space team)
[2] GIPSA-Lab (Control Optimization PErception Navigation and Intelligent Computing team)

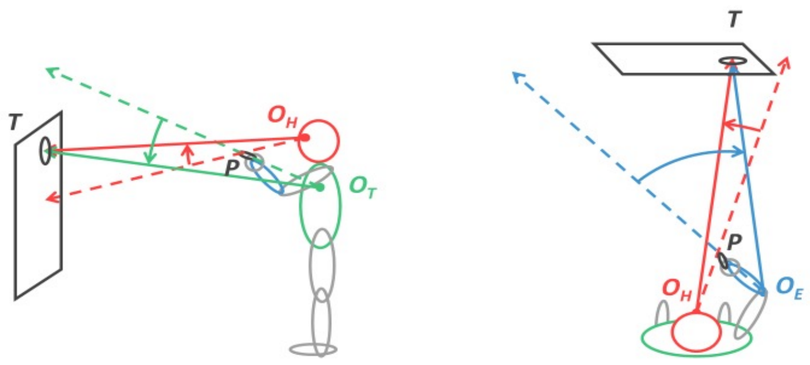
Introduction:

Pointing is a universally used and recognized language, being adopted in childhood before talking. Our goal is to reach a quantitative understanding of pointing, a code implicitly agreed between a pointer and an observer. Spontaneous pointing has been described qualitatively by linguists and semioticians in open situations; the geometry of pointing has been investigated quantitatively mainly in VR and constrained lab conditions. The latter approach generally hypothesizes a direction that originates (*O*) from the eyes and passes through the index fingertip (*P*) to reach the target *(T),* which is not confirmed by the observation of artwork and real people photos. The two approaches may focus of different types of pointing (narrative, loose, sharp). The challenge is to capture the geometric code in any mode of pointing.

1. Pointer experiment

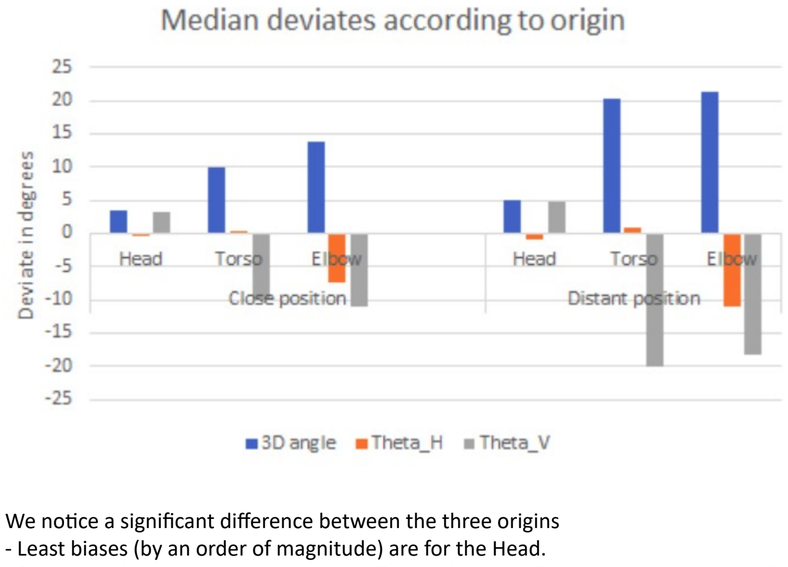
128 postures (two pointer's standing positions and 64 different targets locations).

For each target pointing posture, we measure the 3-D angle *POT* and its projections Theta_V and Theta_H, for each putative origin.



The putative origin O_H (head), O_T (torso) or O_E (elbow) that is closest to real directional origin will give the smallest angle *POT*. Red lines for O_H. *POT* projection on a vertical plane *Theta_V* will be positive if the defined direction *OP* is below the target *T*, i.e. the fingertip *P* is too low, the candidate origin is too high (or the real origin sits lower). *POT* projection on a horizontal plane *Theta_H* will be negative if *OP* is right to the target *T*, i.e. the fingertip *P* is too far right, the candidate origin *O* is too far left (or the real origin sits to the right of the candidate origin).

Results

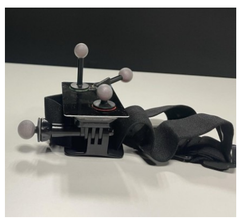


We notice a significant difference between the three origins
- Least biases (by an order of magnitude) are for the Head.
- Strongest bias due to an azimuth bias (Theta_H) for the Elbow, depends on laterality
- However the biases for the eye are not zero (p<001)

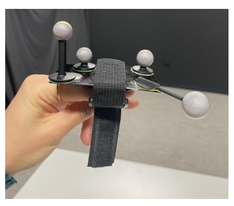
Method Participants are asked to point at targets distributed on a 8x8 vertical panel in a motion capture space to record positions of the fingertip *P* and of three putative origins: O_H (head), O_T (torso) and O_E (elbow).

VICON camera 1

VICON camera 2



Head tracker



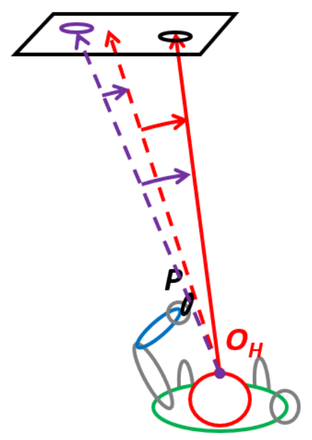
Index tracker

Torso tracker

2. Observer-Pointer experiment

72 situations (permuting three standing positions between pointer and observer for 24 targets T_{Ptr}).

In each target pointing situation, the Observer declares which target T'_{Obs} has been pointed at.

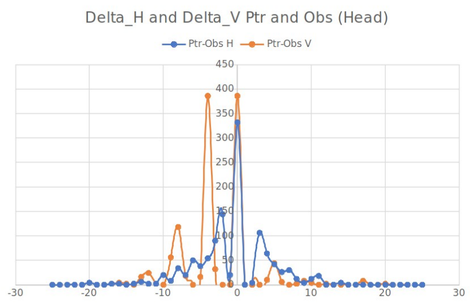


Observer (Obs)

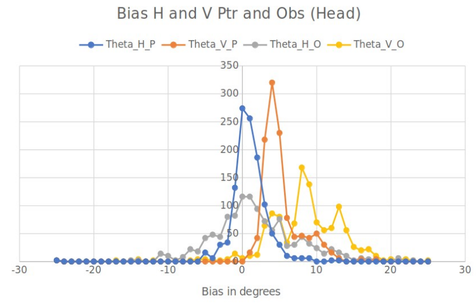
Pointer (Ptr)

The projected Horizontal angles with O_H (head) as putative origin of the pointing direction. The observer (Obs) notes which target T'_{Obs} has been pointed by the Pointer (Ptr) according to their own judgement.
In the example given here, Obs stands to the right of Ptr, and the latter considers the pointed target T'_{Obs} to the right and above the intended target T_{Ptr}. If O_H is the origin of the direction used in Obs' judgement, angle $POHT'_{Obs}$ should be close to zero. The disagreement between POT_{Ptr} and POT'_{Obs} may be evaluated from angle $T'_{Obs}O_HT_{Ptr}$. If both agree on the code and therefore on the target, $T'_{Obs} = T_{Ptr}$, therefore angle $T'_{Obs}OHT_{Ptr}$ tends to zero.

Results



Strong agreement between the pointer and observer on the horizontal axis. On the vertical axis, the agreement takes place only in 50% of the trials.



The vertical component Theta_V, starting from the Head is smaller than the one starting from the torso or elbow origins. However, it is not zero (as is the case for the horizontal component Theta_H) neither for the pointer, nor the observer. Thus, in a directional logic, the aiming hypotheses (ie. an axis having as origin the head) is excluded.

Sharp vs Loose Pointing Experiment

Method

Previous experiments show that cumbersome marker objects (esp. elbow and index) have a strong influence on the participant's way of gesturing (non-natural, 'laboratory-induced' pointing).
We improved the index marker object and removed the elbow object.
We added multiple target sizes and shapes within a larger spatial distribution.
Twofold goal :
- to observe loose and sharp pointing gestures (previous experiments employed only sharp pointing)
- to gather data within an extended domain (used for predictive model training)

Discussion:

Our experiments have confirmed the existence of two pointing methods (sharp and loose).
The use of tracking objects may alter the way of pointing.
Pointing is most natural and effective when employed alongside verbal cues (eg. 'to the right', 'the small one', 'between those two').

Future work:

We propose a natural dialog experiment, eliminating any tracking devices. Participants will be immersed in a collaborative scenario.
A quantitative analysis will be made using video recording + image processing + machine learning (segmentation, clustering, classification) for gesture classification

We propose the development of a Cobot(Collaborative robot) capable of understanding/replicating human pointing (powered by a neural network model that predicts the pointing gesture given a position and a target).