This study\textsuperscript{1} is part of a large-scale cross-cultural research project on the development of spatial language and cognition, in India, Indonesia and Nepal, that focuses on a culturally particular way of organizing small-scale, table space, using a large-scale geocentric spatial orientation system (Dasen & Mishra, in preparation). One of the main questions is at what age this geocentric frame of reference starts to be effective. The study of language development does not provide a clear answer, because young children (ages 4 to 7) use ambiguous “deictic” descriptions, i.e., they just say “this way” accompanied by a gesture. Can these gestures be used to clarify the meaning of language? To answer this question, 234 video recordings of Nepalese children performing the “Perspectives” task (in which they have to describe the location of three objects placed on a table in front of them, under three different conditions) were analyzed separately for both language and gestures. The results show a good correspondence between language and gestures in 9 to 12 year olds. This allows us to interpret further the frame of reference used by the younger children. Out of 367 items on which young children (4 to 9 years) give an ambiguous deictic answer, only 17\% are accompanied with an egocentric gesture, and 83\% with a geocentric one (combining 48\% large gestures linked to the use of cardinal directions, and 35\% medium-large gestures linked to the use of situational local landmarks). This shows that a geocentric frame is at play as early as age 4, even when the child cannot express it clearly in the language.

The research reported in this chapter is part of a larger project on spatial language, encoding and concept development in Indonesia, India and Nepal, focusing on the development of a geocentric frame of reference (FoR) (Dasen & Mishra, in preparation). This is a cognitive process that is unknown in Western developmental psychology, representing a culturally particular developmental path in spatial language and concept development. Using a geocentric FoR means locating objects using a large scale orientation system (such as cardinal directions) even for table-top space inside a room.

One main question of our research is how early children are able to use such a geocentric FoR, either in language (when describing a spatial array) or in other cognitive tasks (such as encoding a spatial array in memory). Much of our data suggests that, in the children we have studied in Bali, India and Nepal (Dasen, Mishra, Niraula & Wassmann, 2006; Mishra, Dasen & Niraula, 2003), a geocentric FoR is predominant, is used very early, and further increases with age. The project reported here seeks further evidence for the precocity of a geocentric FoR in Nepalese children.

\textsuperscript{1} This research was supported by grant 1113-067178 of the Swiss National Science Foundation. We thank Ms. Rena Shrestha and M. Purushottam Tandon for help in collecting the data in Kathmandu under the supervision of Prof. S. Niraula.
Some of the results of a previous study in India and Nepal (Mishra, Dasen & Niraula, 2003) on language development are illustrated in Figure 1. We found that children were starting to use geocentric (G) language by age 6 to 8, with a sharp increase after age 9. Younger children (especially 4 to 5 years, and to some extent up to 8 years) were often using what we have called “deictic”\(^2\) language (D) in describing a spatial array: They just say: “This way/that way”, accompanied by a gesture of a finger or the whole hand.

![Graph showing the percentage of G and D language use by age group in the three locations.](image)

**Figure 1.** Mean proportions of G and D language use on combined tasks, by age group in the three locations (*source: Table 4, p. 375, in Mishra et al., 2003*).

The status of the D category is inherently ambiguous in terms of geocentric vs. egocentric encoding: It could be body related since a body movement is involved, and could mean “to the right/left” (i.e., egocentric), but the movement may also point to a direction that is outside of the display. The latter interpretation is reinforced by the fact that, in both samples in India, the very young children also use what we have called “situational landmarks” (SL) references (i.e., outside of the display, but within the room), a category that is still quite strong at age 6-8, but then disappears with age. At the time, we concluded that the status of D “needs more detailed research, because it could hide markedly different processes (egocentric and geocentric).” (Mishra et al., 2003, p. 379).

Indeed, given the early predominance of geocentric language, D may well be geocentric too (confirming a process that does not exist in Western contexts), but it could also be egocentric, confirming rather the classical (Western) developmental theories claiming that spatial development always starts egocentrically with reference to the body (Miller & Johnson-Laird, 1976; Piaget & Inhelder, 1956).

How, then, could we decide whether D is geocentric or egocentric? Possibly through gestures. Indeed, Levinson (2003) makes a distinction between absolute (geocentric) and relative (egocentric) gestures, in which he describes typical geocentric gestures as being large, made with the outstretched arm or even the whole body, and egocentric ones much smaller, close to the body:

---

\(^2\) Deictic is a Greek word, stemming from “δεικνύω” meaning “show”. Following this, “deictic” means “someone or something showing the way”.

---
[Geocentric] gestures are large, with arms outstretched, allowing accurate sighting of the angles depicted. Unlike our [Western] gestures, they occur in all directions, for example behind the body. (p. 244). … Absolute [geocentric] gestures are large and expansive. [They] are made with fully extended arms, and thus lie naturally far away from the trunk. This contrasts with the data … for American English speakers (said to be the same for speakers of other European languages), where the vast majority of gestures occur right in front of the trunk (Levinson, 2003, p. 252).

Hence, if we were to look more closely at which gestures accompany D language, we might be able to tell which FoR they are in fact using, a geocentric or an egocentric one. What we propose to do is to look first at the language data for the new study in Kathmandu, demonstrating that the D category also occurs there in young children. We then look at the gestures of those whose language is not ambiguous (mainly the older children), to see whether what we think are geocentric and egocentric gestures do correspond to the respective frames in language use. In other words, this will validate our coding for gestures in this particular cultural context. Once we are satisfied with the correspondence between language and gestures, we apply the coding to the younger children, those who use the ambiguous D language. The gestures they predominantly use will tell us in which FoR, egocentric or geocentric, they are functioning.

**Procedure**

The overall sample of this study consists of 400 children aged 4 to 12, in two types of schools (English-medium and Nepali-medium) in Kathmandu, Nepal. One of the language elicitation situations used in our research is the so-called “perspectives” task, in which the children have to describe the location of three non-fronted objects placed on a square piece of cardboard. They give a description from one position, then move to the opposite side of the table and describe it again, and finally the display is rotated by 180°, and the children describe it again. Hence, each child produces 9 items of spatial language (and, possibly, gestures). The situation is illustrated in Figure 2 (the photograph was actually taken during our research in Bali).

The rather cumbersome procedure of moving around the display, or rotating the latter, should not concern us here. It is designed to test the claim that the use of a geocentric FoR potentially allows the same description of a display independently of the position of the speaker. The results of this feature of the study will be reported elsewhere. The display, and the table or desk on which it was placed, were oriented along the cardinal directions, the child initially facing North.

About half-way into the study, we started taking videos of this situation. This means that, although we tested 400 children in all, video recordings are available for 234 of the children (200 of which were from the Nepali-medium schools).

The coding categories for language and gestures are presented in Table 1. In further analyses, we will ignore the intrinsic category, which is irrelevant to our hypothesis (and not very frequent in occurrence).
Table 1. Scoring of language and gestures

<table>
<thead>
<tr>
<th>Language</th>
<th>Gestures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code</td>
<td>Category</td>
</tr>
<tr>
<td>G</td>
<td>Geocentric</td>
</tr>
<tr>
<td>SL</td>
<td>Situational landmark</td>
</tr>
<tr>
<td>E</td>
<td>Egocentric</td>
</tr>
<tr>
<td>I</td>
<td>Intrinsic</td>
</tr>
</tbody>
</table>

The coding of the language was done in Kathmandu, while the coding of gestures was done independently in Geneva, by researchers who did not have the information on the language codes, and did not understand Nepali. Although they could hear the sound recording (if only to be sure of which position was being asked about), they scored gestures independently of any language. Only the first, spontaneous gesture was scored for each position. Out of a potential total of 2,106 items that were scored independently for both language and gestures, 1,583 items were codable for both.

A set of 70 subjects was scored initially for training purposes, and then rescoring when the procedure had been fully established. A fairly high inter-scorer reliability (88%) in scoring gestures was attained after extensive training and discussion.

Results

Language use, complete Kathmandu sample

In our current research project, we again find the same pattern of language use as in our previous studies (note that our previous sample in Nepal was in a rural setting). Figure 3 shows the language used on the Perspectives task.

Figure 3. Language used on the Perspectives task in Kathmandu main study (N=400).

Figure 3 is a plot of the language categories of interest (there are others, such as intrinsic, but they are not of present concern) over age, reporting the mean number of responses in each category out of a maximum of nine. Geocentric language (G) obviously increases with age; in
this case, it corresponds to the use of cardinal directions. Some egocentric language (E: left, right) does occur, but is rather infrequent (note that this is true even in English-medium schools, a finding that will be reported elsewhere). The deictic answers (D) occur predominantly in younger children (4 to 6 years), who also use some situational landmarks (SL: to the door, to the window).

**Consistency between language and gestures**

We first looked at the consistency between language and gestures on the 1,135 items where the language was explicit (non D). The data are presented in Table 2.

**Table 2. Consistent and inconsistent items between language and gestures**

<table>
<thead>
<tr>
<th>Language</th>
<th>S</th>
<th>M</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>66</td>
<td>58</td>
<td>71</td>
</tr>
<tr>
<td>SL</td>
<td>13</td>
<td>64</td>
<td>115</td>
</tr>
<tr>
<td>G</td>
<td>62</td>
<td>332</td>
<td>354</td>
</tr>
</tbody>
</table>

Inconsistency was found on 129 items where E language was accompanied with an M or L gesture and 75 items where SL and G language went with a S gesture, amounting to 18% of the items. Perfect consistency (ES, SLM, GL) was found in 484 items (43%), and we can consider the 447 SLL and GM items to be consistent as well, since they both refer to some direction outside of the display, and at least non-egocentric gestures. The category GM, combining geocentric language with a medium size and short duration gesture, was quite frequent. Total consistency was hence estimated at 82%. This gives us sufficient confidence to conclude that gestures indeed reflect the same frame of spatial reference as language. Our argument is therefore that, since gestures are a good reflection of the frame of reference expressed in language in the older children, they can also be used to determine what frame the younger children are using when they give an ambiguous D verbal answer.

Looking at which gestures go with the Deictic verbal answers (Table 3), only 17% of the answers are accompanied by an egocentric gesture (DS), 35% by the medium scale geocentric gesture (DM) and the majority (48%) by a clear, large-scale geocentric gesture (DL).

**Table 3. Deictic language and three types of gestures**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS</td>
<td>61</td>
<td>17</td>
</tr>
<tr>
<td>DM</td>
<td>130</td>
<td>35</td>
</tr>
<tr>
<td>DL</td>
<td>176</td>
<td>48</td>
</tr>
<tr>
<td>Total</td>
<td>367</td>
<td>100</td>
</tr>
</tbody>
</table>

It could still be the case, if the theory of a universal egocentric developmental starting point was true, that the very young children would use more DS than the older ones. That this is not the case is illustrated in Figure 4.
Even at age 4, Ds accompanied by large geocentric gestures (DL) form the majority, followed by medium geocentric gestures (DM). The mean number of DS is significantly lower than DG at age 4 ($t = 2.384, p < .05$) and significantly lower ($p < .01$) than DL + DG in each age group.

Comparison of English- and Nepali-medium schools

Does learning English in pre-school and primary school relate to the younger children’s spatial frame of reference when they use an ambiguous deictic description? To answer this question, we matched (for age and school grade) the 34 children from the English-medium school for whom video recordings are available with 34 children from the Nepali-medium schools. The English-medium pupils in this sub-sample produced 3 DS items, the Nepali-medium pupils 9. This shows that learning English seems unrelated to a more egocentric FoR.

Conclusion

We have demonstrated in this study that we can rely on gestures to clarify which frames of spatial reference children use when they give ambiguous answers such as “It is this way”. In a situation such as Kathmandu (and indeed in the other locations of our main study, in rural Nepal, in India, in Bali), where both the egocentric and the geocentric frames are potentially available in the language, it could have been expected that young children start with the egocentric reference and move to the geocentric one later. This is clearly not the case. These results come to reinforce our overall findings of a predominance of the geocentric frame in these locations, and add the important conclusion that this indeed starts very early in life. Even at age 4, geocentric gestures predominate over egocentric ones. Whether we would find the same results with even younger children, at ages 2 or 3 years, is a question open for future research –although it would require a lot of patience and the development of specific methods, because it is already very difficult to carry out this sort of testing with four year olds.

In any case, what we are witnessing is a particular developmental path that is different from the one described by mainstream developmental psychology for Western children. If both an egocentric and a geocentric FoR are available in the language, the choice of which one will
be predominant for spatial description and encoding is akin to a cognitive style: both processes are potentially available, but one is preferred over the other. Broad ecological and cultural factors at the macro-system level, as well as more specific ones such as task demands at the micro-system level, will determine the probability of choosing one over the other, in this case the geocentric one.

The precise nature that this geocentric frame of reference takes in these very young children is a matter of debate, since they give no other signs of using this frame in other aspects of non-linguistic cognition (such as in memory encoding tasks). It is possible that these young children produce the large scale gestures because they try to imitate the dominant cultural model they observe in adults, without fully understanding the cardinal orientation system that lies behind it. Nevertheless this no doubt helps them in building up this understanding as they grow older.

This research together with that of Le Guen (this volume) point to the seminal interest of the research by Levinson (2003) and his team into the geocentric frame of spatial reference that is foreign to Western languages and cultures. However the outcomes of both studies show the limitations of any interpretation in terms of extreme linguistic relativism. In both cases, a geocentric frame is used in gestural deixis without the support of geocentric language. Hence we claim that it is cognition that eventually guides language and not the opposite.

References