

HABITUAL PERSPECTIVE

Willem J.M. Levelt

Perspective taking

All four papers presented in this symposium deal in some way or another with the core notion of *perspective taking*. Hence, let me step back for a moment and discuss the role of perspective taking in mediating between language and the rest of cognition.

There are various codes in which the mind can communicate with itself. We have rich representational systems for dealing with visual and tactile space, locomotion, kinesthesia, rhythm, music, social relations, emotions and what have you. We can often directly move from one code to another, e.g. from the visual to the tactile, dependent on the requirements of the task at hand. There is no reason to suppose that there is a single dedicated system, a single language of thought, that is there to mediate among them.

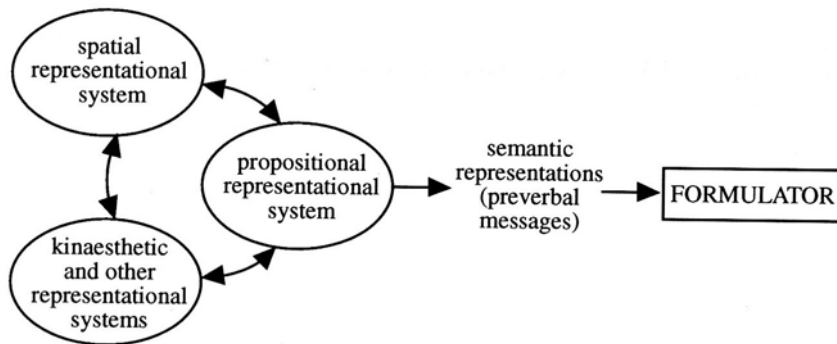


Figure 1. The mind harbors multiple representational systems that can mutually interact. But to formulate any representation linguistically requires its translation into a semantic, “propositional” code (reproduced from Levelt, 1989).

However, as soon as any of these specific information types is to be expressed in language, it must be translated into propositional format (Fig. 1). This is because the semantic representations required to trigger verbal formulation are cast in this format. This requires an act of abstraction, which can usefully be called *perspective taking*. For instance, when talking about a visual scene, we attend to entities that are relevant to the communicative task at hand, and generate predications about these entities that capture them and their relevant spatial relations.

The spatial array in Fig. 2, for instance, can be described in myriad ways, dependent on the perspective you take, such as:

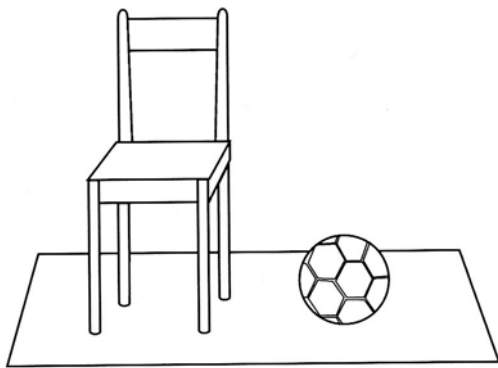


Figure 2. This spatial array can be described from many perspectives.

- (1) a ball and a chair to the left of it
- or
- (2) a chair and a ball to the right of it
- or, equally veridically:
- (3) a chair and a ball to the left of it
- etcetera.

Meanwhile, the notion of perspective taking has been generalized from this limited spatial domain to include any propositional construals of a state of affairs. As early as 1980, Landau and Gleitman discussed the child's perspective problem in referring to the same creature out there as an object, a mammal, a cat, or Felix. These are different lexical concepts that can be mustered to refer to the same entity, dependent on communicative intentions¹.

Perspective taking is free but not arbitrary

Fifteen years ago I presented a paper under the title "Perspective is free" (Levelt, 1990), which captured the view that there is no single, necessary, or hard-wired way of mapping visual, or any other types of mental representation, onto propositional formats for linguistic expression. This does not mean, however, that the choice of perspective is arbitrary. On the contrary. We prefer to relate figures to grounds and not reversely: *the cat on the mat*, not *the mat under the cat*. We go for basic level terms in preference to hyperonyms or hypernoms. We go for spatial perspectives shared in our speech community (relative, intrinsic, absolute), and so on. See Levelt (1996) for detailed discussion. At the bottom of all this is, no doubt, communicative effectiveness.

Perspective taking is also non-arbitrary in a typological sense. If you happen to speak a tensed language, you must, time and again, for your sentence to be grammatical, encode the deictic and intrinsic temporal properties of a state or event, even if this is of no communicative value (as I discussed at length in *Speaking*). In English, expressing deictic proximity (such as in *here/there* or *this/that*) requires a bipartite conceptual

¹ For an excellent review of perspective taking in this broader sense and the child's acquisition thereof, see Eve Clark (1999).

distinction between proximal/distal to ego. For Spanish or Japanese speakers, however, the required conceptual distinction is tripartite, proximal/medial/distal. When you are a speaker of English you will have to think of number (one or more) any time you are going to use a count noun. Languages differ in interesting ways in the features they mark obligatorily. Almost any lexical domain involves language-specific distinctions. One further well-studied case is the domain of topological spatial relations, such as expressed in English *in*, *at*, *on*. No two languages carve up this domain in the same way, as Bowerman and Choi (2001) have shown. A speaker must, time and again, acknowledge the appropriate distinctions. I have argued (in *Speaking*) that this language-specific obligatory perspective taking, which Dan Slobin has usefully called “thinking-for-speaking”, is fully routinized in adult speakers. Attending to the relevant features has become habitual.

Comments on Papafragou’s paper

Anna Papafragou’s paper provided such evidence for still another domain, motion verbs. For English speakers this lexical domain typically requires the encoding of manner, whether communicatively interesting or not. Not so for Greek speakers, whose verbs of motion typically do not express manner, but only path. Papafragou then argued that the Greeks, different from the speakers of English, will only consider to encode manner if it is not already pragmatically implicated in the event. And that is exactly what she found.

Comments on Landau & Dessalegn’s paper

Linguistically encoding a feature, whether obligatorily or not, can deeply affect your mental bookkeeping. Barbara Landau’s paper (with Dessalegn) shows that, by linguistically attending a subject to the spatial location of a color patch in a geometrical figure, the subject’s performance in recognizing that figure improves dramatically. *How* this works is another issue. It can indeed be the case, as argued in the paper, that language instructs the visual system to bind the features of color and location, such improving visual recall. But alternatively, the subjects may have done no more than keeping the linguistic encoding, *red on the bottom*, in memory, using it to good effect in the recall task. I would like to see what happens if the experiment is performed while the subject is shadowing speech, which will interfere with memorizing such a linguistic code.

Does thinking-for-speaking spill over into non-linguistic cognition?

Let me now return to the obligatory encoding. It has become a hot issue whether the routine, habitual thinking-for-speaking a language requires, spills over into non-linguistic thought. This is, of course, an empirical issue, to be decided by careful experimental analysis. A-priori, however, it is entirely reasonable to entertain the hypothesis that linguistic experience indeed provides us with new concepts, applicable beyond the strict use of language. I vividly remember Landau and Gleitman (1980) first reporting on the blind child’s very early acquisition of the notion SEE. The following research program, involving Susan Fisher and others, made it abundantly clear that many concepts are initially acquired as lexical concepts through syntactic bootstrapping, not through non-linguistic experience. It is unlikely, in my opinion, that this richness of new, linguistically induced concepts is only and exclusively used for linguistic encoding.

Levinson et al.'s frames of spatial reference paradigm

The present test bed for Whorfian spill-over is the experimental paradigm developed by Steve Levinson and colleagues in our Max Planck Institute. The essence of the paradigm is to test whether speakers of a language are inclined to use in non-linguistic tasks the spatial frame (or frames) of reference encoded in their language. Applying these frames of reference in linguistic encoding is a pure case of thinking-for-speaking. Levinson's group has surprisingly shown that there are just three basic types of spatial perspective used in the large variety of language communities they studied, called intrinsic, relative and absolute perspective. Any language uses one, two or all three of them and none of them is universally used. English uses relative and intrinsic perspective, Guugu Jimithir uses just absolute perspective, Mopan, a Mayan language, uses almost exclusively intrinsic perspective. There are no obvious communicatively relevant ecological or other non-linguistic factors that determine the choice of perspective. Whatever its cause, the perspective has become standard routine, habitual thinking-for-speaking.

In their experiments, Levinson and colleagues have shown marked preferences for using their language-dependent perspectives in non-linguistic tasks (see Levinson, 2003, Majid et al., 2004 and Levinson and Wilkins, in press, for overviews). The main task involved the subject memorizing a linear arrangement of three toy animals (out of four) and then, after some delay, reconstructing the array from memory, after 180 degrees rotation to another table. The reconstruction largely follows absolute or relative perspective, dependent on the subject's language. There are other remarkable non-linguistic correlations with linguistic perspective. Users of absolute perspective, who must always dead reckon to keep track of their absolute bearing (such as North-South), turn out to have refined directional sense, incomparable to the sloppy directional sense of relative perspective users. In relating events, users of absolute perspective, different from users of relative perspective, align their gestures with directional veridicality.

These findings have been challenged by Li & Gleitman (2002) and Li, Abarbanell & Papafragou (2005) and we saw some of that in Lila Gleitman's paper at this symposium. Both she and I lack the time to sketch the whole picture of these controversial findings. I have decided to limit myself to two points which have, I would like to say, unnecessarily complicated the discussion.

Ability or preference?

The first point is this. Much confusion resulted from not carefully distinguishing between *ability* and *preference*. The discussion should be about preference, about habitual thought. Ability cannot be the issue. Just as we can learn a language that uses a different perspective from our own, we can learn to use that perspective anyhow if necessary for dealing with some non-linguistic task. Initial texts have indeed been confusing here².

² Pederson et al. (1998): "Far more than developing simple habituation, use of the linguistic system, we suggest, actually **forces** the speaker to make computations he or she might otherwise not make".

Gumperz & Levinson (1996): "Now, a speaker of such a language **cannot** remember arrays of objects in the same way as you and I, in terms of their relative location from a particular viewing angle..."

Gleitman & Papafragou (in press): "... different communities of humans, speaking different languages, **would think differently** in just the extent that languages differ from one another.

If you still want to test ability, as Li, Abarbanell and Papafragou (2005) did, you must give your subjects just one option and see whether they succeed. That is correctly argued. Hence, Li et al. forced speakers of an absolute language (Tzeltal) to use, under different conditions, relative or absolute perspective, or at least they thought they did. Subjects turned out to be equally proficient under both conditions. Regrettably, the authors did not notice that their relative (or “egocentric”) condition can also be solved under absolute perspective, and this is also the case for their swivel chair experiment³. So, after all, they did *not* force a choice and therefore cannot draw any conclusions about their subjects’ egocentric or geocentric abilities.

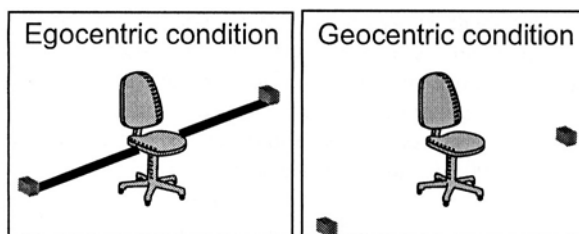
Moreover, a mere thought experiment would have sufficed to refrain altogether from the swivel chair ability enterprise. Take a Guugu Jimitther (i.e., exclusively “absolute”) subject, ask him to put his house key in his trouser pocket. Then turn him around blindfolded. Then ask him to retrieve the key. If he desperately starts exploring both pockets, he is a Whorfian victim. Would you really want to do such an experiment?

Landmarks as nuisance variable

The second unhappy complication resulted from confusing the use of absolute perspective with the use of landmarks. That happened in the Li and Gleitman (2002) paper, which claimed to test a preference for absolute responses in subjects by setting up controlled landmark cues, the paper pond with kissing ducks, also used in the experiment presented here. English speaking subjects now showed a preference for aligning the toy figures towards the landmark, which was taken as evidence that they happily applied absolute perspective. Levinson et al (2002) replied that the duckpond landmark could not have functioned as an absolute bearing. Absolute systems are abstract. Visible landmarks do not play an essential role in their use. The duckpond wasn’t even fixed in space: it moved from one table to another. It was simply part of the scene to be reproduced, essentially inviting intrinsic perspective. Distinguishing preference for absolute versus landmark-oriented (or intrinsic) requires a 90 degree rotation, Levinson et al claimed. That experiment they performed with Dutch subjects (who don’t have an absolute frame)

Ibidem: “It would be quite amazing if, among all the creatures that walk, fly and crawl on earth, only humans in virtue of acquiring a particular language **lose the ability** to use absolute and relative spatial coordinate frameworks flexibly”. [all bolding mine]

³ Consider this case of a shivel chair trial: An “absolute” subject in the shivel chair “egocentric condition”, will start observing that he is oriented to the North and that the coin is put in the box to the West. After blindfolded turning around, the subject will think that he is now oriented East (notice that it is totally irrelevant whether this is veridical) and infer that, therefore, the coin must be in the North box.



The authors are apparently also unaware of the fact that Tzeltal has extensive body part terminology, including terms for left/right hand.

and the result was corresponding: they happily produced landmark-oriented (i.e. intrinsic) responses, but no absolute ones.

In her reported experiment Lila Gleitman now tested the 90 degree condition on Tzeltal and Tsotsil subjects. The former have both absolute and intrinsic perspective, the latter have relative and intrinsic perspective. Both populations dominantly produce landmark-oriented responses. Both do have the relevant intrinsic frame available, so no surprise they can give an intrinsic response. But why don't the Tzeltal also happily give absolute responses and the Tsotsil relative responses? My hunch is that the landmark condition has deeply confused the experimentation. In order to give the Figure 3(a) response (as subjects apparently do) and not the Fig. 3(b) response (as they do not do), but which is a landmark-oriented response, there must be a reason for the subjects to align the array with the table front or with their egocentric left-right dimension (as already argued by Levinson, 2003). They are probably guessing that the experimenter reasonably wants them to do such aligning in instructing to “make the array the same”.

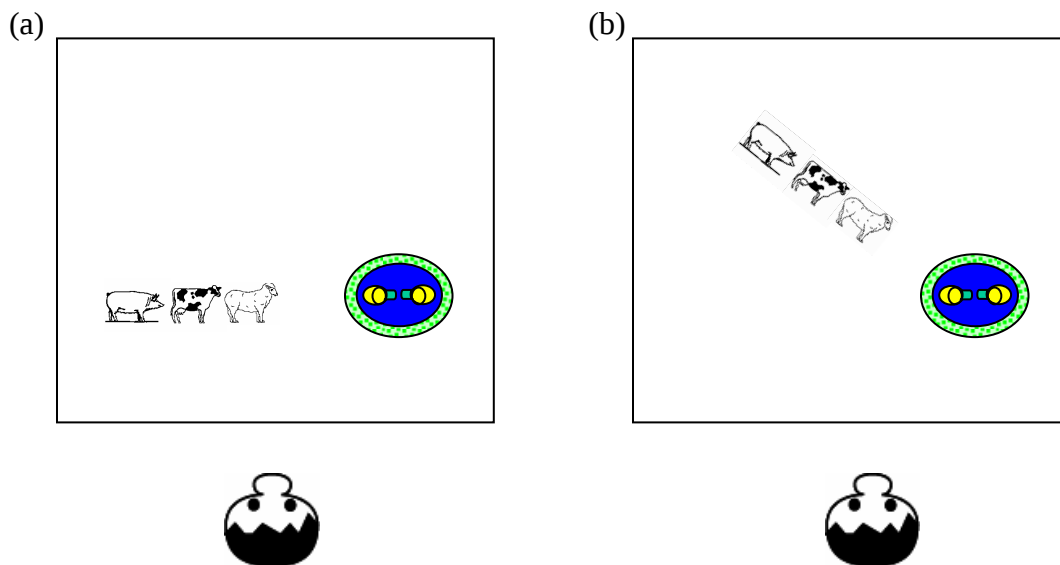


Figure 3. The 90° rotation condition. Subjects align “towards the landmark” as in (a), not as in (b); they find it desirable to align with the table side or with their own left-right dimension. Notice also that the landmark itself has an intrinsic orientation.

Moreover, the duckpond landmark is not neutral, but has its own intrinsic orientation, the “kissing-direction” of the two ducks. Hence, there is more involved than just applying intrinsic perspective. We are studying the ruminations of subjects about what they think the experimenter wants them to do in a highly ambiguous situation. As Michotte said to me in 1958 when he was 77 and I was 20: “the art of experimentation is largely knowing what experiments *not* to do”. Let us not unnecessarily introduce nuisance variables, such as oriented landmarks, when we want to compare habitual, unreflected preferences for absolute and relative frames in non-linguistic memory tasks. This will only keep confusing the discussion.

REFERENCES

- Bowerman, M. & Choi, S. (2001) Shaping meanings for language: universal and language-specific in the acquisition of spatial semantic categories. In: M. Bowerman & S.C. Levinson, *Language acquisition and conceptual development*. Cambridge: Cambridge University Press.
- Clark, E. V. (1997) Conceptual perspective and lexical choice in acquisition. *Cognition* 64, 1-37.
- Gumperz, J.J. & Levinson, S.C. (1996), Introduction to Part 1. In Gumperz, J. & Levinson, S. (Eds.), *Rethinking linguistic relativity*. Cambridge: Cambridge University Press.
- Gleitman, L. & Papafragou, A. (in press) Language and thought. In: K. Holyoak & B. Morrison (Eds.), *Cambridge handbook of thinking and reasoning*. Cambridge: Cambridge University Press.
- Landau, B. & Gleitman, L. (1980?) The language of perception in a blind child.
- Levelt, W.J.M. (1989) *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.
- Levelt, W.J.M. (1990) Perspective is free. MM-Symposium MIT
- Levelt, W.J.M. (1996) Perspective taking and ellipsis in spatial descriptions. In P. Bloom, M.A. Peterson, M.F. Garrett, & L. Nadel (Eds.), *Language and space* (pp.77-107). Cambridge, MA: MIT Press.
- Levinson, S.C. (2003) *Space in language and cognition. Explorations in linguistic diversity*. Cambridge: Cambridge University Press.
- Levinson, S.C. & Wilkins.D. (in press) *Grammars of space*. Cambridge: University Press.
- Levinson, S.C., Kita, S., Haun, D.B.M., & Rasch, B.H. (2002). Returning the tables: language affects spatial reasoning. *Cognition*, 84, 155-188.
- Li, P., Abarbanell, L., & Papafragou, A. (2005) Spatial reasoning skills in Tenejapan Mayans. First poster session, this conference.
- Li, P. & Gleitman, L. (2002) Turning the tables: Language and spatial reasoning. *Cognition*, 83, 265-294.
- Majid, A., Bowerman, M., Kita, S., Haun, D., Levinson, S. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Science* 8(3), 108-114.
- Pederson, E., Danziger, E., Wilkins, D., Levinson, S., Kita, S & Senft, G (1996). Semantic typology and spatial conceptualization. *Language*, 74, 557-589.
- Slobin, D. (1996) From 'thought and language' to 'thinking for speaking'. In: J. Gumperz & S.C. Levinson (eds.), *Rethinking linguistic relativity*, '70-'96. Cambridge: Cambridge University Press.