POTENTIAL STAGES IN THE CULTURAL EVOLUTION OF SPATIAL LANGUAGE

MICHAEL SPRANGER

Sony CSL Paris, 6, rue Amyot,
Paris, 75005, France
spranger@csl.sony.fr

This paper investigates the function and formation of spatial language involving landmarks. We argue for two points. First, landmarks have an important function in spatial language. Agents that can use different landmarks and express their choice perform better than agents who cannot. Second, the use of landmarks has a positive effect on evolving systems of spatial relations. Agents that have the cognitive ability for using landmarks and agents that can express their choice of landmark are more successful in building spatial relation systems than those without. Based on this evidence the paper hypothesizes potential stages in the evolution of spatial language. Communicative success and in combination with environmental conditions can be driving forces for agents to go from a stage without landmarks to systems that allow arbitrary objects to be used as landmarks.

1. Introduction

A primary function of spatial language is to locate objects in spatial environments. Here is an example of a *locative* phrase from German.

(1)	*der*	*Block*	*links*	*der*	*Kiste*
	the.NOM	block.NOM	left.PREP	the.GEN	box.GEN
	'The block to the left of the box'.				

This seemingly simple phrase conveys a particular way of conceptualizing reality by integrating a number of syntactic and semantic *strategies*. The phrase combines the *projective* spatial relation "links" (left) with a landmark – an inanimate object denoted by the phrase "der Kiste" (the box). The speaker choose to explicitly *mark* the landmark (not obligatory). The relationship between the spatial category and the landmark is conveyed using the German case system (obligatory).

The strategic choices manifest in such an utterance are culturally coordinated. The fact that inanimate objects can serve as landmarks for spatial relations is true for many languages but certainly not for all. Some languages restrict landmark use to only animate objects or only humans, or – even more radical – just to trees (Levinson, 2003). Moreover, the strategies for expressing conceptual choices – lexicon and grammar – differ across languages (Svorou, 1994).

The diversity of strategies is something that comes as no surprise to people studying spatial language across cultures (Levinson & Wilkins, 2006; Levinson, 2003). But many questions remain about the evolutionary mechanics of spatial language. Recently, a number of attempts are made at explaining the evolution of aspects of spatial language such as toponyms (Schulz et al., 2006), the role of perspective – "me" and "you" – (Steels & Loetzsch, 2008), the origins of conceptualizatian strategies (Spranger, 2011b) and grammar (Spranger et al., 2010). In this paper we focus on the role of landmarks including perspective but also inanimate objects. We follow standard methodology for evolutionary explanations (Tinbergen, 1963) applied to the cultural level (Steels, 2006).

2. Spatial Language Games

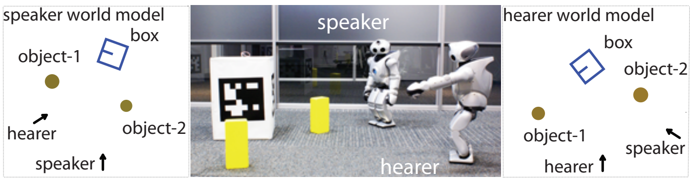

Figure 1. This figure shows the set-up for spatial language game experiments. The images left and right show the internal situation model as perceived by each robot.

We use *language games* to study the cultural origins of spatial language. Two robots from a population are interacting in a shared environment and are trying to draw each others attention to objects in the vicinity using language. The set-up is exemplarily shown in Figure 1. The environment consists of a number of blocks of equal size and color (circles), a box (rectangle) and the interlocutors (arrows). The vision system of each robot tracks objects in the vicinity and establishes a model of the environment consisting of blocks (circles) with real-valued distances and orientations of objects with respect to the body of the robot. The environment is open-ended. New blocks, boxes and robots are added or removed and their spatial configuration changed.

1. Each agent perceives the environment after both have established a joint attentional frame (Tomasello, 1995).

2. The speaker randomly selects an object from the situation model (further called the topic T). The speaker tries to find a discriminating spatial relation applied to a particular landmark or perspective for describing T. Subsequently, the speaker looks up the words he associates with the spatial relation and the landmark and produces an utterance.

3. The hearer looks up which relation is associated with the utterance and also tries to look up words for landmarks in memory. He examines the context to find out whether there is a unique object which satisfies the spatial relation with respect to the landmark. The hearer then points to this object.

4. The speaker checks whether the hearer points to T. If the hearer pointed correctly the game is a success and the speaker signals this outcome to the hearer. If the game is a failure, the speaker points to the topic T.

There is a lot to say about processing machinery necessary for making these games work. For this paper, we have developed techniques and operational models that allow agents to flexibly and autonomously conceptualize spatial reality and express these conceptualizations in situated interactions (Spranger, 2011a).

3. The Function of Landmarks in Spatial Language

We can use language games to quantify the function of landmarks in spatial language with respect to communicative success. For this we equip agents with different language and conceptualization strategies and measure their performance in multi-agent simulations. In all populations agents are using *projective* spatial relations equivalent to English "front", "back", "left" and "right". The relations are defined as a similarity functions which are computed from the difference in angles between an object and the prototypical angle of the category (Spranger, 2011a). We compare 5 populations.

egocentric (ego) Agents use a single spatial relation in each utterance which is always interpreted egocentric to each robot. For instance, if the speaker conceptualizes "left" to refer to an object, the hearer will interpret the relation from his perspective.

perspective, unmarked (pp, um) Agents use the perspective of themselves but also of the other robot (similar to Steels & Loetzsch, 2008). Upon hearing the term "left" the hearer retrieves the best possible interpretation taking into consideration his own and the other robot's perspective on the environment. Similarly, the speaker might use an utterance such as "left" but refer to a position left of the hearer rather than left of himself.

perspective, marked (pp, m) Agents are using different perspectives but also clearly marking which perspective was used. For example, the speaker might say "left you" to indicate spatial relation and perspective.

landmarks, unmarked (lm, um) This population can use perspective (robots) and allocentric landmarks (the box), in conceptualizing reality. They do not communicate which object was used.

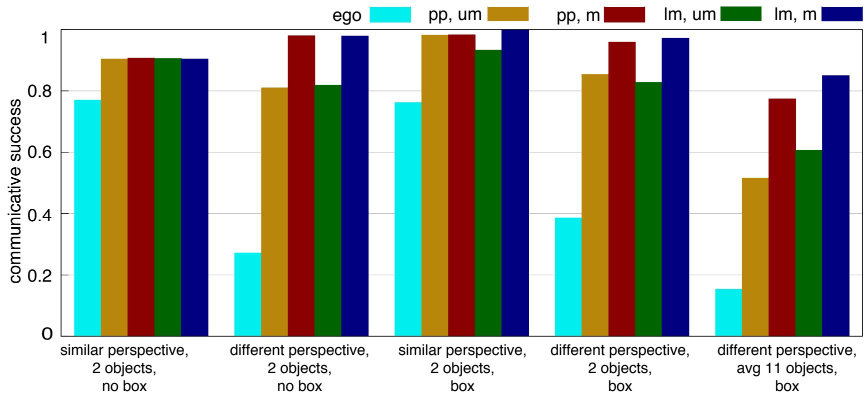

Figure 2. Results for populations of 10 agents tested in different experimental conditions. For each result agents interacted 2000 times. Every successful interaction counts with 1.0, otherwise with 0.0. These scores are averaged over all interactions.

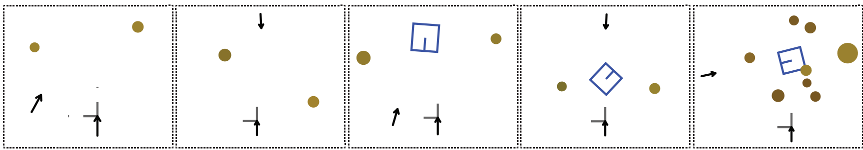

Figure 3. Examples of spatial scenes from each experimental condition ordered from left to rights as in Figure 2. Left, for instance, agents (black arrows) share a similar perspective on the spatial scenes, there are always 2 objects per scene (circles) and no box. To the right the most difficult condition is shown. Agents have differing perspectives, there are on average 11 objects and there is a box.

landmarks, marked (lm, m) Same population as *landmarks, unmarked*. But, agents express which landmark they use in all utterances.

Figure 2 shows communicative performance of each population in different experimental conditions. Each condition consists of many spatial situations, each of which include two robots and blocks whose positions change. But, the conditions differ in whether a box is present or the average number of objects and the typical perspective of robots (see Figure 3).

Clearly, the results show that agents capable of using different perspectives and landmarks perform better in communication than agents without. The effect increases the more objects are in each spatial context and the more the viewpoints of the interlocutors on the scene diverge. This is because agents which use different landmarks are much more flexible. They have more choice of relation-landmark pairs and often can choose one that fits best or better for discrimination. There are three important observations. First, if the environment is simple and agents have a similar perspective on the scene, it is sufficient to use egocentric interpretation of phrases. Second, the best strategy is to use any landmark possible (both robots and allocentric landmarks). Third, marking which landmark or perspective was used always outperforms the unmarked case in which agents use different landmarks but cannot express them.

The *unmarked* populations are interesting because they effectively show possible and maybe necessary intermediate evolutionary steps between a purely egocentric spatial language strategy and one that allows landmarks to be used and expressed. To develop a full-blown landmark systems, agents have to go through an unmarked stage where using a particular object as landmark is possible.

4. The Formation of Spatial Relation Systems and Landmarks

Previous work has shown how agents can culturally evolve lexical systems of spatial relations (Spranger, 2011a). Here, we apply these insights and study the effect of landmarks on such systems. We re-use the populations introduced in the previous section. All spatial relations (and terms) are removed from agents and instead they are provided with the following language change operators,

Invention Speakers invent new spatial terms and categories when they cannot discriminate a topic T from the context.

Adoption Upon hearing an unknown term the hearer uses the information from the speaker's pointing to learn not only the term but also the spatial relation the speaker might have used.

Alignment Invention and adoption happen in single interactions and it takes time for the new relations and terms to spread. Alignment operators orchestrate the gradual coordination of linguistic knowledge in the population. Agents continuously update their internal representations (Garrod & Doherty, 1994) by changing category representations, i.e. prototypical angles, over time and by rewarding and punishing categories and words based on success in interactions.

Figure 4 shows the dynamics of these operators over time. Agents start without spatial relations and spatial terms. Gradually they invent and align their linguistic repositories. At the same time the population becomes more successful in communication. Figure 5 compares different populations in a number of experimental conditions (same as used earlier). When agents do not have the means to use landmarks they fail in developing successful systems of spatial relations (if the environment is complex). Again the populations which can express the landmark they are using are outperforming agents which cannot. Agents which can use the full scope of landmarks (me/you/box) available in scenes outperform those which can only use perspective (me/you) which outperform those that can only use an egocentric strategy.

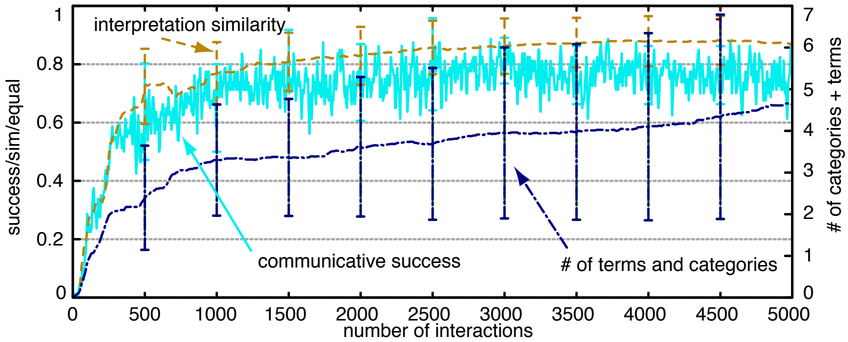

Figure 4. Dynamics of a population of 10 agents which develop a spatial category system (egocentric strategy) of 5 spatial relations in the *similar pp, 2 obj, no box* condition (5000 interactions, 8 trials). At the same time the spatial relations align across the population (see rise in *interpretation similarity*).

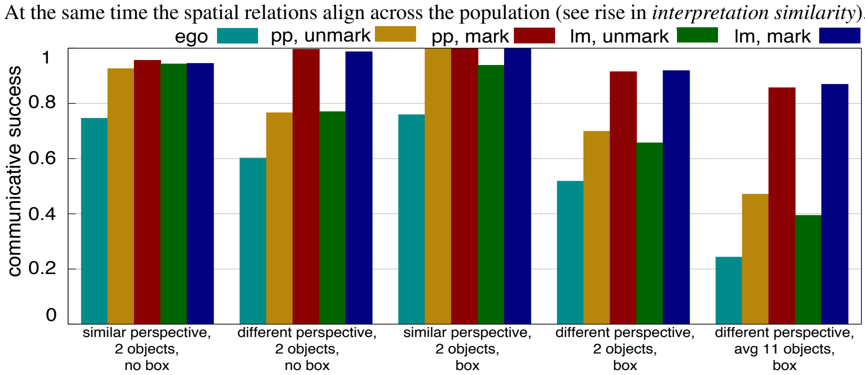

Figure 5. Comparison of different populations forming spatial relation systems (same experimental conditions as in Figure 2, 8 trials). Every population consists of 10 agents which first develop a spatial language system (5000 games) and then their communicative success is measured (2000 interactions).

Notice, how *landmark, unmarked* performs worse than the *perspective, unmarked*. The reason is that, upon hearing a new term agents have to adopt it using a particular landmark. This decision is essentially random, because agents were given no other means of deciding. The *landmark, unmarked* population has more landmarks to choose from. Therefore, agents are more likely to choose wrong. This makes clear that intermediary stages are necessary and that perspective strategies are suitable intermediaries.

Lastly, Figure 6 shows that agents can evolve a system of spatial relations while *at the same time* evolving a lexicon for expressing which landmark was used. The following is an example utterance one agent in such a population uttered

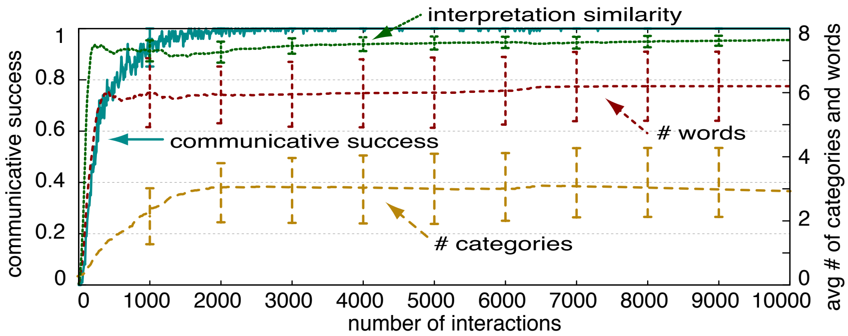

Figure 6. This population forms a system of projective spatial relations and words for marking perspective and landmarks. There are three the difference between the average number of words (*# constructions*) and the average number of spatial relations (*# categories*).

after 5000 interactions.The agent is using a category that he developed which behaves similarly to "left".

(2) *wutobe* *taketo*
box category-12
'to the left of the box'.

5. Discussion and Outlook

The experimental evidence presented in this paper suggests two important points. First, the function of landmarks in spatial language can be defined in terms of communicative success. Landmarks enable agents to communicate successfully about spatially distributed objects. Second, necessary stages for the evolution of spatial language such as unmarked conceptualization strategies are possible. It is obvious that before a landmark can be expressed or marked in language it must have a particular function in conceptualization. The question is whether such an initially unmarked case does lead to higher success so that the intermediate stage is sustainable. In the experiments, unmarked populations are performing better than the egocentric-only populations which shows that they are sustainable. Furthermore, marking of landmarks outperforms both the unmarked and the egocentric case. Judging from the results, perspective is another possible intermediary between egocentric systems and full landmark systems.

Communicative success is not the only area where use of landmarks can have positive effects. Steels and Loetzsch (2008), for instance, argue for the impact of perspective marking on cognitive effort which is the amount of processing required on the side of the hearer. The same argument also holds for the systems presented in this paper but it was omitted for space reasons.

Spatial language does not evolve in a vacuum. Rather the interaction of strategies described in this paper with other strategies such as toponyms, but also the interaction of different syntactic strategies (see Levinson & Wilkins, 2006 for an overview of different strategies) are important for understanding why particular strategies are present in one language but not in others. These are issues which certainly merit exploration and remain to be studied.

Acknowledgements

The author thanks Martin Loetzsch and Luc Steels for help in gathering the robotic data. Funding was provided by Sony CSL Paris and by the EU FP7 project Alear.

References

Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition*, *53*(3), 181–215.

Levinson, S. C. (2003). *Space in Language and Cognition: Explorations in Cognitive Diversity.* Cambridge University Press.

Levinson, S. C., & Wilkins, D. (2006). *Grammars of Space.* Cambridge University Press.

Schulz, R., Stockwell, P., Wakabayashi, M., & Wiles, J. (2006). Towards a spatial language for mobile robots. In *The evolution of language (evolang 6)* (pp. 291–298). World Scientific.

Spranger, M. (2011a). *The evolution of grounded spatial language.* Unpublished doctoral dissertation, Vrije Universiteit Brussels (VUB), Brussels, Belgium.

Spranger, M. (2011b). Recruitment, Selection and Alignment of Spatial Language Strategies. In *Advances in artificial life, ecal 2011.* MIT Press.

Spranger, M., Pauw, S., & Loetzsch, M. (2010). Open-ended semantics co-evolving with spatial language. In *The evolution of language (evolang 8)* (pp. 297–304). World Scientific.

Steels, L. (2006). How to do experiments in artificial language evolution and why. In *The evolution of language (evolang 6)* (pp. 323–332). World Scientific.

Steels, L., & Loetzsch, M. (2008). Perspective alignment in spatial language. In *Spatial language and dialogue* (pp. 70–88). Oxford University Press.

Svorou, S. (1994). *The grammar of space* (Vol. 25). John Benjamins.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, *20*(4), 410–433.

Tomasello, M. (1995). Joint attention as social cognition. In *Joint attention: Its origins and role in development.* Lawrence Erlbaum Associates.