Two Predicted Universals in the Semantics of Space
Author(s): Terry Regier
Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Semantic Typology and Semantic Universals (1993), pp. 479-489

Two Predicted Universals in the Semantics of Space

Terry Regier
University of California at Berkeley

1 Introduction

This paper presents two predicted linguistic universals concerning the semantic structuring of space. These posited universals are predicted by a neurally and psychologically inspired connectionist model of the acquisition of lexical semantics in a visually-grounded domain (Regier 1992). They are presented here by way of opening them to empirical falsification.

While the predictions are the central focus of this paper, there are two related issues which arise, and which are also discussed here.

The first of these is that the work presented here suggests by example that the construction of detailed computational models of the *acquisition* of lexical semantics may be a productive avenue for the pursuit of semantic inquiry generally. In particular, this work suggests that models of this sort can give rise to falsifiable predictions regarding semantic universals, based on what the model may say concerning the *learnability* of particular combinations of semantic features.

The second of these concerns the issue of learning in the absence of explicit negative evidence. The model presented here learns using a technique which relies on positive evidence only.

We proceed to examine first the predictions themselves, then the model of semantic acquisition which gives rise to them, including the means by which the model learns without explicit negative evidence. Finally, we provide an indication of just how it is that the model gives rise to the predictions.

2 The Predictions

The two predictions are the *endpoint configuration prediction* and the *endpoint polysemy prediction*. These are presented in turn below.

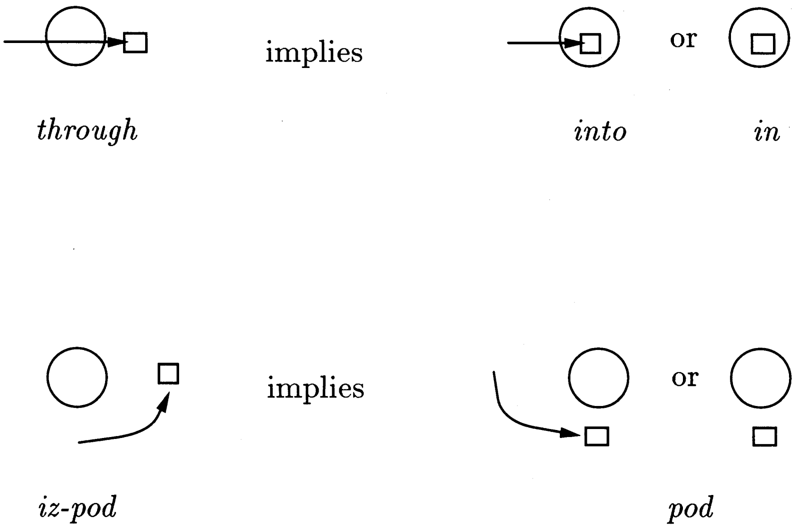


Figure 1: The Endpoint Configuration Prediction: Any language which has a lexeme denoting motion out of or motion through some configuration will also have a lexeme denoting either motion *into* that configuration or static location in it.

2.1 The Endpoint Configuration Prediction

The endpoint configuration prediction, illustrated in Figure 1, states that

> Any language which has a lexeme denoting motion out of or motion through some configuration will also have a lexeme denoting either motion *into* that configuration or static location in it.

Consider, for example, the English lexeme *through*, shown in the upper left of Figure 1. Here we see a small square trajector[1] moving through a circular landmark object. We can view the lexeme *through* as denoting motion through a configuration of inclusion, in that inclusion of the trajector within the landmark occurs in mid-event for *through*. This leads us to predict, by our posited universal, that English will have a lexeme denoting either motion

[1]The *trajector* is an (often moving) object located relative to another object, known as the *landmark*.

into a configuration of inclusion (*into*), or static location in the configuration of inclusion (*in*).

Consider the lower example, from Russian. The fact that Russian has a lexeme *iz-pod*, denoting motion out from underneath, i.e. motion of a trajector out from underneath a landmark, leads us to predict that Russian will also have a lexeme denoting either motion into the region under the landmark, or static location under the landmark. In fact Russian *pod* can be used in either of these two senses.

2.2 The Endpoint Polysemy Prediction

The endpoint polysemy prediction, illustrated in Figure 2, states that

> The use of a single lexeme to denote either static location in some configuration or motion into that configuration will be more likely to appear in a language than the use of a single lexeme to denote either static location in a configuration or motion *out of* that configuration.

Thus, this prediction states that lexemes such as English or German *in*, which can denote either static location inside or motion into the region inside, are more likely than are lexemes like the fictitious spatial term *in/out-of*, which denotes either static location inside or motion out of the region inside.

Similarly, lexemes such as Russian *pod*, meaning either location underneath or motion to the region underneath, are more likely than are lexemes like the fictitious Russian *pod/iz-pod*, meaning either location underneath or motion out from underneath.

3 The Model

Having presented the predictions themselves, we now move on to the model that gives rise to them, beginning with a more precise specification of exactly what function it performs: the model learns to linguistically classify visually perceived events. It is given a set of movie-lexeme pairs, and learns to label the simple events and relations shown in the movies with the appropriate lexemes. For example, consider Figure 3. This figure is a representation of a movie of the sort the model is exposed to during training. Each movie, like this one, contains two objects, a *landmark*, or object relative to which other objects are located, and a *trajector*, an often moving object located relative to the landmark. In this figure we see a movie of five frames, with the frames overlaid one on top of another. Thus, the five circles are in reality a single circular trajector moving out from underneath a horizontally-extended

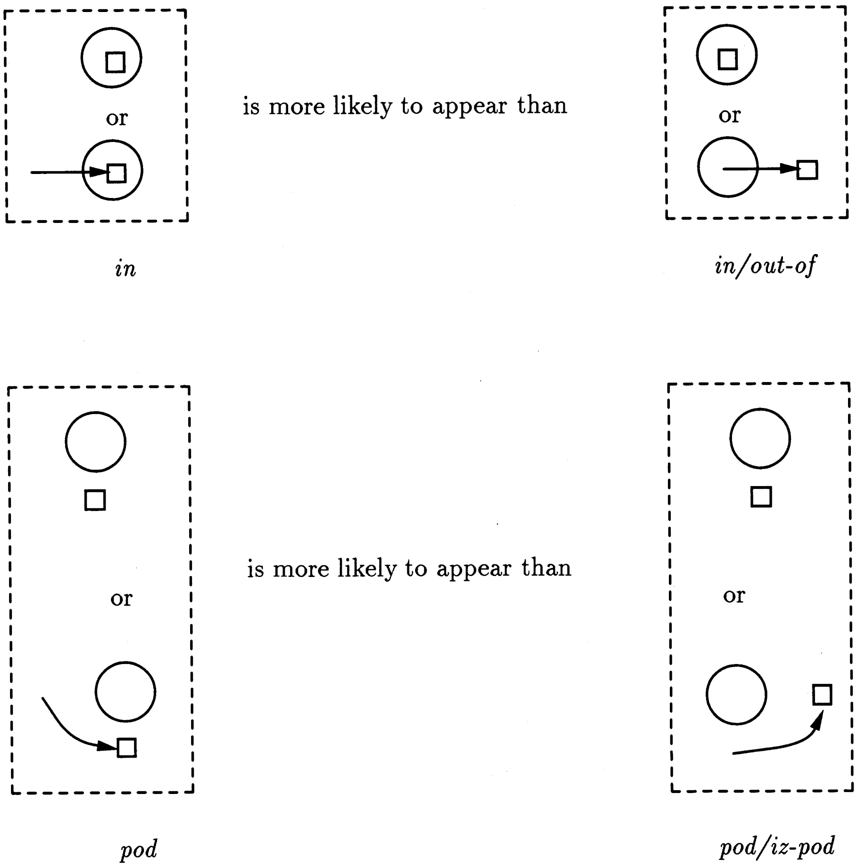


Figure 2: The Endpoint Polysemy Prediction: The use of a single lexeme to denote either static location in some configuration or motion into that configuration will be more likely to appear in a language than the use of a single lexeme to denote either static location in a configuration or motion *out of* that configuration.

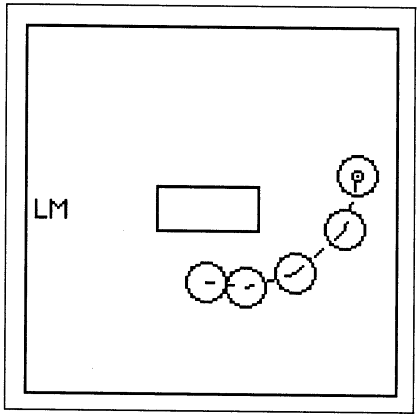


Figure 3: A trajector moving relative to a landmark. The final position is indicated by a small dot inside the trajector. This event illustrates the Russian preposition *iz-pod* (out from underneath).

rectangular landmark, in five steps. The final frame of the movie is indicated by a small dot inside the trajector. Note that the landmark, labeled "LM", remains stationary throughout this movie.

This movie illustrates the Russian preposition *iz-pod*, meaning "out from underneath", and was in fact labeled with that preposition as part of a Russian training set. After training, the model was able to label other movies showing motion out-from-underneath as *iz-pod.*

The model was not built specifically to lead to the predictions we are discussing here. Rather, it was built to address two overarching linguistic issues.

The first issue is that languages differ dramatically in the ways in which they structure space. The model is intended to be able to learn the spatial system of any language, and therefore to be able to shed light on universal human processes of semantic acquisition in the domain of space. It has so far learned spatial terms from Bengali, English, German, Mixtec, and Russian.

The second issue concerns learning language without the benefit of explicit negative evidence, as children appear to (Braine 1971; Bowerman 1983; Pinker 1989). The model presented here similarly learns without the benefit of explicit negative evidence. This is done through the use of *deliberately weakened implicit negative evidence*, a somewhat bulky name for what is in reality a very simple notion, and a notion which may be taken as a solution to the problem of no negative evidence generally. The beginning of the idea, in its specific implementation in this domain, is to take a positive instance of one

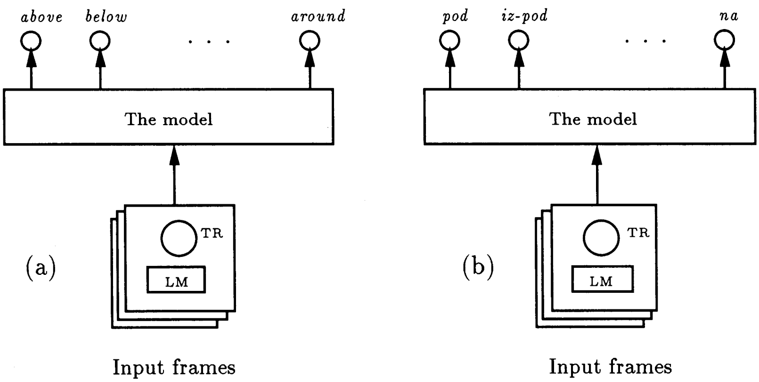
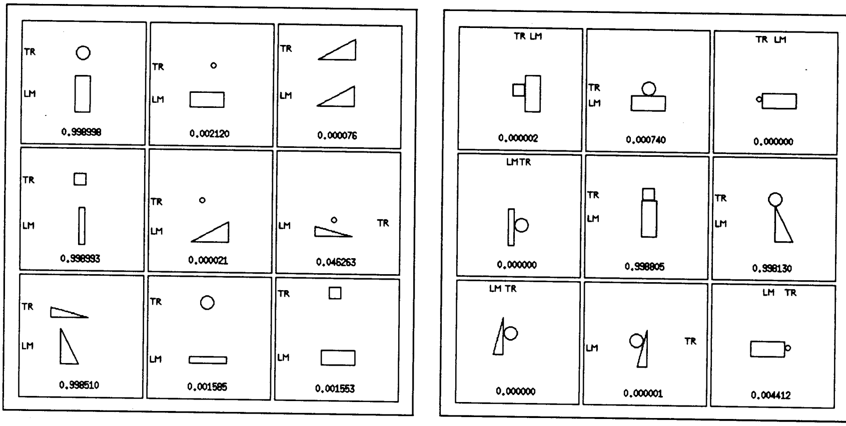


Figure 4: Model configurations for English and Russian

spatial term as an implicit negative instance of all others. This is the "principle of mutual exclusivity" (Markman 1987). For example, if a particular spatial relation is labeled *above*, we assume that that also means "not *below*", "not *inside*", and so on. Unfortunately, this heuristic also leads one to conclude that a positive instance of *above* is a negative instance of *outside*, which it is not. In fact, every positive instance of *above* is also a *positive* instance of *outside*, which means that the heuristic is fallible. However, we can salvage this fallible heuristic by deliberately weakening the evidence from implicit negatives. Intuitively speaking, this means taking the implicit negative evidence less seriously than explicit positive evidence. Using this simple approach, the model is able to learn without exploiting explicit negative evidence.

Figure 4 presents the model, configured to learn a set of English spatial terms in (a), and a set of Russian spatial terms in (b). Note that the model learns a *system* of spatial terms in consort; this is critical if the mutual exclusivity heuristic is to be used in obtaining implicit negative evidence from positive evidence only.

The input to the model is a movie of the sort shown in Figure 3, and the model is trained so that when a movie portraying some event is shown, only those output nodes corresponding to lexemes which accurately describe the event are activated. For example, if the movie in Figure 3 were supplied to the model shown in Figure 4(b) after training, the *iz-pod* output node would become activated, indicating that the model has classified the movie as a positive example of *iz-pod*.



(a) (b)

Figure 5: Mixtec *šini* tested on positive examples of English *above* and *on*. *Šini*, meaning "head", indicates that the trajector is either above or on top of a vertically extended landmark.

The model learns to determine, for each language to which it is exposed, which spatial features – such as inclusion, contact, alignment with some axis, and the like – are semantically significant. It learns to linguistically classify visually perceived events by noting when in the event significant spatial features appear. A detailed specification of the model, including further explication of the technique of deliberately weakened implicit negative evidence, may be found in (Regier 1992).

4 Results

No claim is made that the current model is in fact capable of learning the spatial system of any language. However, the model can learn the spatial systems of a range of languages, some of which differ quite significantly from English in their structuring of space. It is thus a first step on the road toward a model which will be capable of fulfilling the original goal.

Consider for example Figure 5, which illustrates the model's performance on the lexeme *šini* from Mixtec, an Otomanguean language spoken in the state of Oaxaca, Mexico. Brugman (1983) presents a semantic analysis of spatial terms in Mixtec, spelling out the manner in which spatial locations are referred to as metaphorical body-parts. *Šini*, which translates literally to "head" in English, can be thought of as indicating that the trajector is either above or

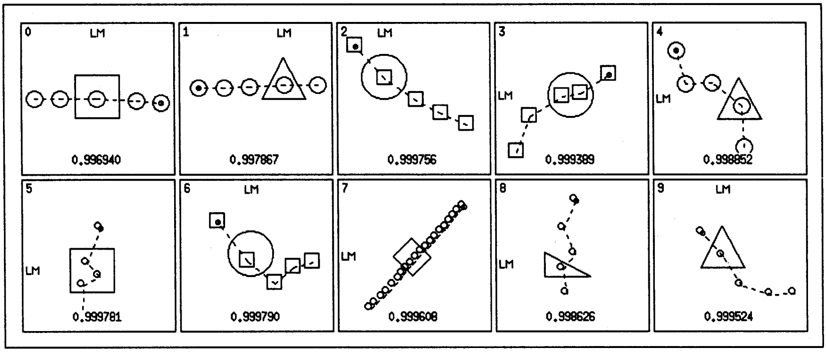


Figure 6: English *through*.

on top of a vertically extended landmark – in a position corresponding to the "head" of an erect biped. It may seem at first that learning this spatial system would require a good amount of world knowledge which the current model does not have access to, e.g. basic anatomy of bipeds. However, it seems to be the case that much if not all of the categorization is based on primitive, purely perceptual features of the landmark object in question, such as the orientation of its major axis, which the current model is able to extract directly from the image.

The figure illustrates the model's performance, after it was trained to respond strongly to instances of Mixtec *šini*, to instances of English *above* and *on*. Responses range from 0.0 if the input is considered an extremely poor example of *šini*, to 1.0 if it is considered an excellent example. Since not all instances of *above* and *on* are considered good examples of *šini*, this figure illustrates an example of cross-linguistic variation in spatial structuring, as well as the model's ability to learn spatial systems which differ radically from that of English.[2]

Figure 6 illustrates the model's ability to linguistically categorize events involving motion. The model was first trained[3] to recognize a set of English spatial prepositions, including *through*, and was then exposed to the movies displayed here, each of which is a positive instance of *through*, and none of which had ever been presented to the model earlier. As indicated by the

[2]The other major goal of the project, learning without negative evidence, is in fact not illustrated by this example, as *šini* was learned using explicit negative evidence. This was done since not enough was known about the language to construct a full contrast set. Once such a set has been constructed, the system will be trained on Mixtec spatial terms without the use of explicit negative evidence.

[3]Training in this example took place without explicit negative evidence.

numbers at the bottom of each movie, all are considered excellent examples of *through* by the model. Analogously, the model gave very low *through* ratings to poor examples of *through.*

These two examples are meant to demonstrate the model's ability to learn linguistic spatial systems. The model was constructed to address the two linguistic issues mentioned earlier: cross-linguistic variation in spatial systems, and the issue of learning without negative evidence. The two linguistic predictions with which this paper is centrally concerned, the endpoint configuration prediction and the endpoint polysemy prediction, in fact resulted as unexpected byproducts of the model-building process.

5 Predictions from the Model

Now that the predictions and the model itself have been presented, we move on to indicate just how it is that the model gives rise to the two predictions.

Recall that the first prediction is the endpoint configuration prediction, and that its essence is that any language which has a lexeme denoting motion out of or motion through some configuration will also have a lexeme denoting either static location in that configuration or motion into it. Thus the existence of *through* in English implies the existence of a lexeme such as *in* or *into.*

This is predicted by the model because of a particular type of endpoint emphasis which the model exhibits: *the model will learn to detect only those spatial features which occur at the end of some movie it has seen.* And since the model is exposed to only positive instances of spatial terms from the language being learned,[4] this means that it will only learn those spatial features which occur either at the endpoint of some named event or in a named static configuration.

Thus, if the model is learning a set of English prepositions and is not exposed to movies for lexemes denoting events (or relations) ending in inclusion, such as *in* or *into*, it will not be able to learn to detect the feature of inclusion, and will therefore not be able to learn *through.* This is so since its failure to detect inclusion will lead to a failure to be able to learn *through*, inclusion being critical to *through.*[5] Thus, the existence of *through* in a language implies the existence of lexemes like *in* or *into.*

This characteristic of the model, the fact that it only learns to detect features which are present at the end of some training movie, was not built in deliberately, but rather resulted from the attempt to build as simple a model as possible. The prediction thus arises from a model that was arrived at not

[4]Recall that the model makes use of deliberately weakened implicit negative evidence, which is derived from positive examples only.

[5]Nor for that matter will it be able to learn any other lexeme the semantics of which involve inclusion.

by specifically trying to embody this phenomenon, but on the independent grounds of structural simplicity.

The endpoint polysemy prediction, on the other hand, is arrived at on somewhat more empirical grounds. Recall that this prediction states that we are more likely to find a polysemous lexeme denoting either location in a configuration or motion *into* that configuration – as in the case of English or German *in* – than we are to find one denoting either location in a configuration or motion *out of* that configuration. This prediction, like the first, results from a form of endpoint emphasis which the model embodies, but one which is most easily characterized in terms of empirically observed learning and generalization behavior, rather than in-principle learnability constraints. The observed behavior is as follows: when the model is trained on English *in* in only its static sense, it often generalizes to consider the motion-into sense a good example of *in* as well, despite never having seen a motion-into movie labeled *in*. However, the model will only rarely generalize from static *in* to motion-out-of, i.e. motion which began in and then moved out. It is based on this and other analogous observed behavior that the endpoint polysemy prediction is proposed.[6]

6 Conclusions

This paper has presented two predicted universals concerning the linguistic categorization of space, and the model which led to them. The predictions are presented with the intention of opening them to empirical falsification.

In addition, this work illustrates a more general point: that the construction of detailed computational models of the *acquisition* of lexical semantics affords a fruitful approach to the study of semantics *per se*. Such models can give rise to falsifiable semantic predictions of the sort presented here, based either on arguments of learnability or on empirically observed generalization tendencies.

Finally, the work presented here also provides a simple solution to the problem of learning language in the absence of explicit negative evidence, through the use of deliberately weakened implicit negative evidence.

[6]It is worth noting that there are examples of polysemous lexemes which can denote either location in a configuration or motion out of that configuration, e.g. Russian *u*, which can mean either "by/beside", or "from". The prediction, however, maintains only that these will be less common than polysemous lexemes with static location and motion-into senses, not that they will not exist. On the other hand, the first prediction, the endpoint configuration prediction, is a strict implicational universal; a single counterexample would falsify it.

References

BOWERMAN, MELISSA. 1983. How do children avoid constructing an overly general grammar in the absence of feedback about what is not a sentence? In *Papers and Reports on Child Language Development*. Stanford University.

BRAINE, M. 1971. On two types of models of the internalization of grammars. In *The Ontogenesis of Grammar*, ed. by D. Slobin. Academic Press.

BRUGMAN, CLAUDIA, 1983. The use of body-part terms as locatives in chalcatongo mixtec. in Report No. 4 of the Survey of California and other Indian Languages, pp. 235-90. University of California, Berkeley.

MARKMAN, ELLEN M. 1987. How children constrain the possible meanings of words. In *Concepts and conceptual development: Ecological and intellectual factors in categorization*. Cambridge University Press.

PINKER, STEVEN. 1989. *Learnability and cognition: The acquisition of argument structure*. MIT Press.

REGIER, TERRY, 1992. *The acquisition of lexical semantics for spatial terms: A connectionist model of perceptual categorization*. Computer Science Division, EECS Department, University of California at Berkeley dissertation. available as Technical Report TR-92-062, International Computer Science Institute, Berkeley.