Ewald Lang (HUB & FAS, Berlin)

Basic Dimension Terms: a first look at universal features and typological variation*

0. Introduction

Though published some 25 years ago, Berlin & Kay's (1969) pioneering study on color terms (which the title of the present paper alludes to) has not found its counterpart in the field of spatial vocabulary up to now. This is indeed a surprising gap in the literature - in view of the observation that all languages seem to have a certain sample of lexical items to make reference to spatial dimensions such as height, length, width, depth, distance etc.[1], and taking into consideration that our spatial concepts more or less directly originate in human perceptual endowment, which provides the source of their supposed universality.

In the past two decades, there was a considerable amount of research work on space perception, shape recognition, visual discrimination etc. which attempted to prove or at least to support the claim that "the dimensions languages pick out are just those dimensions the human perceptual apparatus is tuned to pick out" (cf. Clark 1973, Clark & Clark 1977). Plausible as this guideline undoubtedly may be, things turned out to be more complicated. In the meantime we know that there is more to dimensional designation than perception-based categorizing of axes, planes, extensions etc. and simple projecting of top - bottom, front - rear, left - right sides from some observer-based body-schema (cf. Herskovits 1986) onto spatial objects.

In a series of studies (notably Bierwisch & Lang (eds.) 1989; Lang 1990 a,b; Lang, Carstensen & Simmons 1991), the grammar of Dimension Assignment (DA), i.e. the set of conditions according to which natural languages pick out and lexically encode spatial dimensions, has been shown to comprise at least the following components:

(I) **Position and gestalt properties.** DA is basically organized not by a single body-schema but by two interacting categorization grids called "Primary Perceptual Space" (PPS) and "Inherent Proportion Schema" (IPS), both being independently traceable to human perceptual endowment ;

(II) **Parameters.** DA makes use of a limited set of Dimension Assignment Parameters (DAP) which - emerging from the grids in (I) - go to make up primary candidates for being lexicalized and thus provide the basis to look for universals of how natural languages encode spatial dimensions;

(III) **Conceptual-Semantic Interfaces .** DA involves a set of devices that account for the differentiation between, as well as the interaction of, non-linguistic conceptual *world knowledge* and language-bound *word knowledge* as regards the way in which spatial objects are assigned primary and/or contextually induced dimensions in terms of linguistic expressions.

---

* Earlier versions of the paper were presented at the FAS Inaugural Conference " Linguistic Universals & Typological Variation", Berlin, March 17-19, 1994, and at the L.A.U.D. Workshop "Language & Space", Duisburg, March 22-24, 1994. Special thanks go to Paul Kay (Berkeley) for encouraging discussions on the methodology of such an investigation. For providing and/or checking the data which this study draws on I am indebted to Byong-Rae Ryu [Korean], Horst D. Gasde, Chen Xuan [Mandarin], Joanna Blaszczak [Polish], Svetlana Poljakova, Vladimir Klimonov [Russian], and Marcela Adamiková [Slovak]. For stimulating ideas I am grateful to Zubin & Choi (1984), Zubin & Svorou (1984).

[1] Bierwisch's seminal paper of 1967, which proposed a handy set of features, inspired much work in acquisition research but has not been challenged as semantic classic or revised until the appearance of Bierwisch & Lang (eds.) 1989.

After having gained some idea of what is going on in DA in general and having worked through the grammar of DA in German and English in some detail, which inter alia includes a PROLOG implementation (Lang, Carstensen & Simmons 1991) that served to prove and improve the DA model outlined in (I) - (III), we now may feel encouraged to tackle questions of universal features and typological variation in the realm of basic dimension terms.

The paper is divided into five sections. Section one presents a brief overview of the framework mentioned in (I) - (III) above; sect. two offers a couple of preliminary universals of lexicalization regarding dimension terms; sect. three illustrates typological variations within the scope of the universals outlined so far; sect. four examines - within a group of cognate languages that share the same set of lexical items for DA - a case of variation that is induced via interference by a language from outside the group; sect. five gives a brief outlook on problems to be tackled in the future.

1. Overview of the framework

In order to be brief, I will outline the analysis of DA as presented in Lang 1990 a,b; Lang, Carstensen & Simmons 1991 by means of a summarizing diagram - cf. Fig. 1 below. It indicates the way in which perceptual information (from vision, organ of equilibrium, upright walk etc.) is being conceptually categorized by two independent but interacting grids, PPS and IPS, each of which contributes in its own way to delimit and identify what is to be taken as a spatial dimension.

(1) The main tenet is that DA to spatial objects works by designating certain axis extensions of a given object as **spatial dimensions** by picking out some axis extension **d** of object **x** due to

(a) **d**'s coincidence with an axis of PPS (e.g., an axis **d** of object **x** is designated as "**x**'s height" if **d** coincides with the Vertical, or as "**x**'s depth" if **d** coincides with the Observer Axis etc.) or/and

(b) **d**'s showing some distinctive gestalt feature as defined by IPS (e.g., an axis **d** of object **x** is designated as "**x**'s length" if **d** is the Maximal axis of **x**, or as "**x**'s thickness" if **d** is the SUBstance Axis of **x** etc.).

Fig.1 below shows three levels on which spatial information involved in DA has to be represented.

The Perceptual Level is determined by sensory perception that emerges from how our senses instatiate physical parameters, in the case of spatial perception it is above all those parameters that in one way or other derive from the force of gravity.

The Conceptual Level is determined by what results from categorizing perceptual input by means of PPS and IPS in view of its relevance to human behaviour. In other words, perceptive distinctions are conceptualized only to the extent that they are needed for the naive physics which underlies our everday knowledge of space. The conceptual categorization of spatial objects by PPS and IPS provides us with an inventory of spatial features that are essential to the way in which DA works in natural languages. The role of this inventory of features is twofold:

(2) The spatial features that emerge from PPS and IPS, while keeping their conceptual content constant, occur in two representational formats:

(a) as entries of so-called Object Schemata (OS), that is, as elements which our conceptual representations of spatial objects are made of (= Conceptual Level in Fig. 1)

(b) as linguistically relevant Dimension Assignment Parameters (DAPs), that is, as semantic elements which the meanings of dimension expressions are made of (= Semantic Level in Fig. 1)

It is important to note that the DAPs form but a designated proper subset of the inventory of spatial features defined by PPS and IPS, that is, only a subset of the entries in OS also occur as DAP. This reflects the basic idea that DA rests on designating certain specified object extensions as spatial dimensions. The components shown in Fig 1. will be briefly commented upon in the sequel.

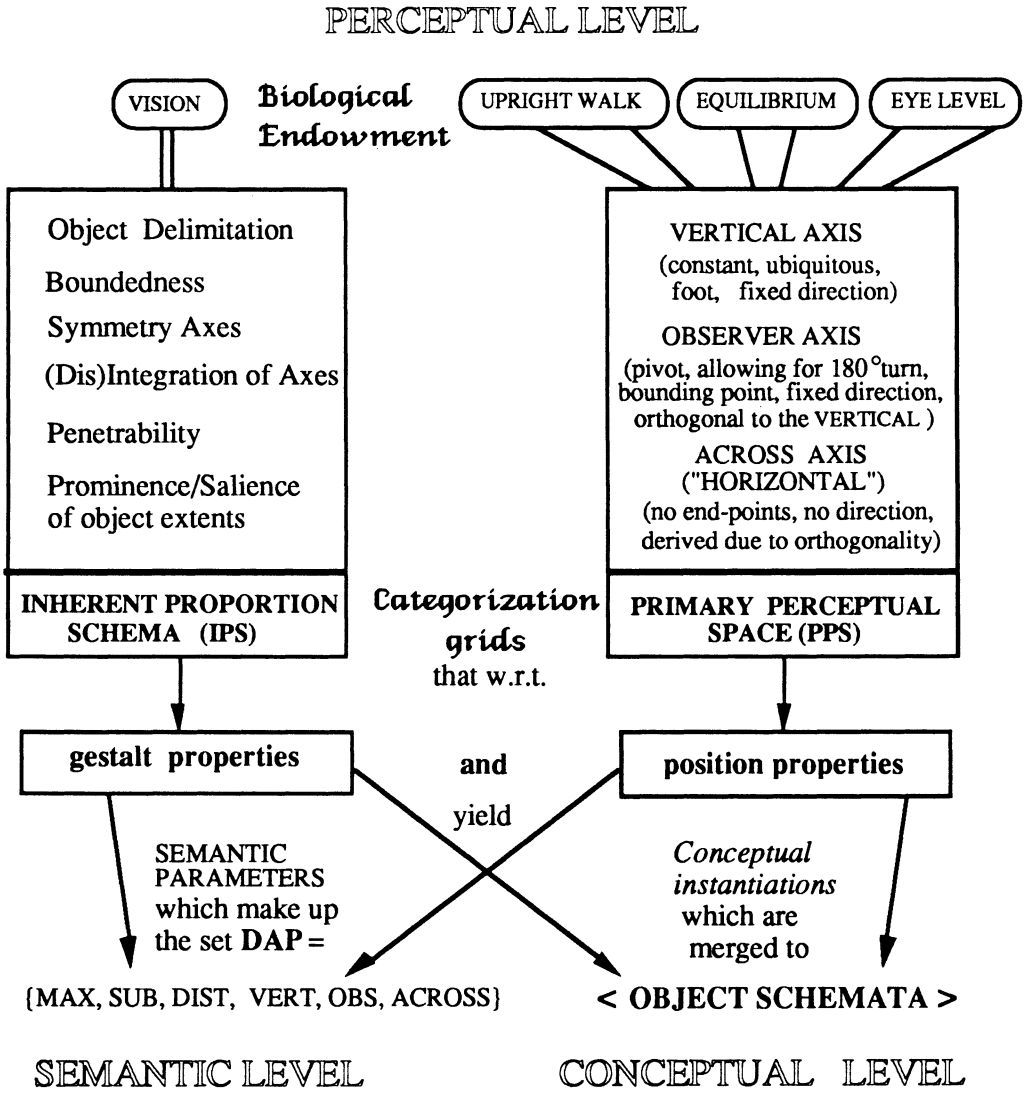


Fig. 1 Cognitive components involved in assigning spatial dimensions to objects

1.1 Primary Perceptual Space (PPS).

PPS consists of a system of axes that form our internal model of the external space. The PPS axes include the Vertical, the Observer axis, and the Horizontal or Across axis, each of which has its own properties relevant to the way in which we conceive of the spatial environment around us.

(3) It is by reference to PPS as a categorization grid that objects
  (a) are assigned a **position** and/or a **location** in space
  (b) are **localized** with respect to one another
  (c) are said to **move** (i.e. to change their position and/or location)

Taking reference to the Vertical as a case in point, we may illustrate (3) (a - c) with respect to some of the spatial relations between the objects shown in **Fig. 2** below by the sample given in (4).

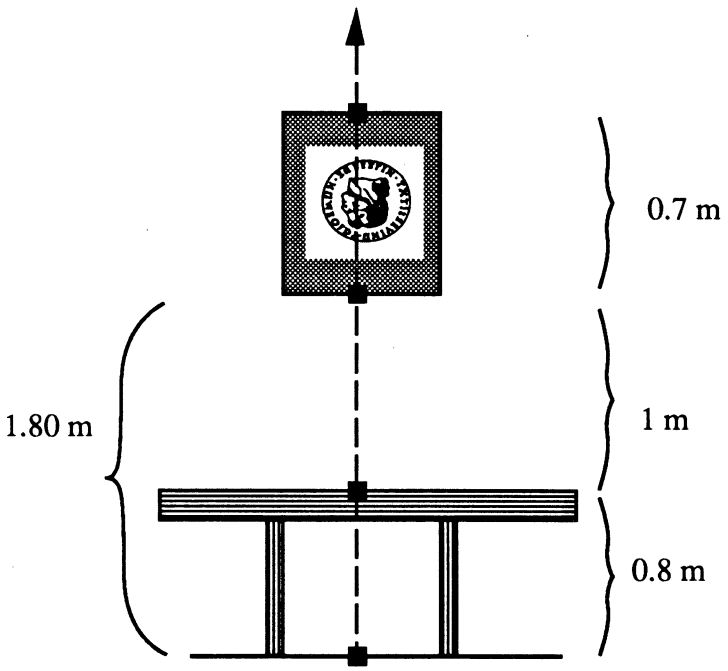


Fig. 2

(4) (a) The *table* is 0.8 m high / 0.8 m in height / is standing / is upright [position]
The *poster* is 0.7 m high /0.7 m in height / is hanging / is upside down
The *poster* is 1.80 m high, at a height of 1.80 [location]

(b) The *table* is under /below the poster [localization]
The *poster* is 1m above / 1m higher than the table

(c) The *poster* is aligned with the table / turned upside down [change of position]
The *poster* was lifted / raised to a place above the table [change of location]

First, all underlined items in (4) involve reference to the Vertical of the surrounding space. This obviously forms a constitutive part of their lexical meaning and is being represented by the parameter VERT on the Semantic Level (cf. Fig 1 and 1.3). Thus, VERT is present in the meaning of each of these items. However, the way VERT is packaged into the lexical meanings differs, depending inter alia on the syntactic category of the given item (for dimension adjectives - cf. Lang 1989; for position verbs - cf. Maienborn 1990a,b, 1993; Lang, Carstensen, Simmons 1991; for projective prepositions - cf. Lang 1991, 1993).

Second, in addition to that, reference to the Vertical is also an essential part of our conceptual knowledge of spatial objects such as "table" or "poster". The concept "table" (and possibly also the semantic representation of the word *table* [2]) includes the feature of having a canonical orientation. A table has an axis **d** which is designated for being aligned to the Vertical and which thereby defines a table's canonic height, i.e. its canonic top - bottom extension, and hence its normal position in the spatial environment. The concept "poster" (and possibly also the semantic representation of all picture nouns) includes the feature of having an inherent orientation. A poster (or picture etc.) has an axis **d** which (due to the object it shows) is assigned an inherent height, i.e. an inherent top - bottom extension, absolutely independent of the Vertical and of the poster's actual position.

When lexical items like *stand, lie, upright, upside down* etc. are predicated of spatial objects, they explicitely refer to such conceptually fixed verticality features. Based on this the examples in (4) get regular interpretations whereas things like **The ball /line/hole is standing upside down* are ruled out as (conceptually) deviant. Hence, having a fixed, canonic or inherent orientation or no orientation at all, is part of the representation of spatial objects on the Conceptual Level (cf. Fig.1). Technically, this is achieved by providing the Object Schemata of "hill", "table", "tree" "poster" etc. with an obligatory primary entry *vert* (which is only typographically distinct from the semantic parameter VERT) while preventing the Object Schemata of "ball", "line", "hole" etc. from being available for any verticality specification whatsoever.

Instead of telling the analogous story about the Observer Axis of the PPS, the semantic parameter OBS involved in the meaning of *deep, in front of, behind* etc. and in the canonical vs. inherent perspectivization of spatial objetcs on the Conceptual Level (e.g. "river", "cupboard" vs. "hole", "wound"), I refer to Lang, Carstensen, Simmons (1991).

I should note, however, that the third axis of the PPS, the so-called Horizontal or, as I prefer to say, the Across axis, is rather different in that it lacks the power of categorizing spatial objects into subclasses based on putative features of canonical or inherent horizontality or transversality. Due to their differing origin in human perceptual endowment, the three axes of the PPS have distinct properties (briefly noted in Fig.1, in more detail discussed in Lang 1989) and hence each of them plays a distinct role in determining our internal model of the external space. In short:

(5) The **Vertical**

> Originating from the effects of gravitation as perceived by the organ of equilibrium, the Vertical is an orientation cue which is **ubiquitous and constant**, that is, available everywhere and with the same effect at all times. The Vertical is physically and conceptually the **most salient** and also the **dominant** axis of PPS ; the other axes are defined in relation to the Vertical.

---

[2] The question of whether all or only a specified subset of the spatial features emerging from PPS and IPS should be considered to be part of the lexical meaning of *table, poster* etc. is one of the facets of the world knowledge vs. word knowledge debate within the two-level approach to semantics (cf. Lang 1994). For the purpose of this paper, I will adopt the latter view without further ado.

(6) The **Observer axis**

Originating in the visual organ, this axis is determined by the line of sight of a (potential or actual) **observer**. Due to this, the Observer axis is flexible in two respects:
(a) it is - unlike the gravitation based Vertical - not anchored in the physical space but rather induced by a movable and moving human interpreter of the physical space;
(b) it has an anatomically determined pivot allowing for a 180° turn in either of two planes.

In the **unmarked** case, given by the position of the eyes of an observer in upright posture, the Observer Axis is orthogonal (at **90°**) to the Vertical. In the **other relevant** configuration, the Vertical and the Observer axis lie at an angle of **180°** such that they run parallel but in diametrically opposed directions. The Observer axis is the source of depth perception.

(7) The **Across axis**

The third axis of PPS is not an axis we are endowed to identify by primary perceptual information, but is **derived from**, hence **dependent on**, the two others just to fill the gap determined by the properties of the latter. The Horizontal or Across axis is exclusively defined by its **orthogonality** to the Vertical and to the Observer Axis.

PPS inter alia provides us with the semantic parameters VERT and OBS (and possibly ACROSS) that are relevant to DA. As we will see in sections 3 and 4, the different status of these axes has direct consequences for delimiting the scope of universal features and typological variation in the lexicalization patterns that cross-linguistically underlie basic dimension terms.

1.2. Inherent Proportion Schema (IPS)

Spatial objects are furthermore categorized by their gestalt properties, i.e. by features drawing on whether or not an object has boundaries, symmetry axes, salient axes (e.g. a Maximal axis, which terms like *long, short, tall, along* etc. refer to), on whether an object axis is visually penetrable (distance axis) or not (substance axis), and on an object's dimensionality (1D, 2D, 3D). IPS provides us - among other things - with the semantic parameters MAX, SUB, DIST (and possibly ACROSS) that are relevant to DA.

1.3 Dimension Assignment Parameters (DAP)

The **interaction** of the two categorization grids PPS and IPS provides us with an inventory of semantic Dimension Assignment Parameters (DAP) and with an inventory of Object Schemata (OS) of conceptually admissible objects specifying the full range of dimension, position, and mobility properties of spatial objects. The latter inventory, which yields a complete catalogus mundi of possible spatial objects, cannot be repeated here (see Lang, Carstensen, Simmons 1991).

I will, however, enumerate the inventory of semantic DAP (with the exception of the holistically assigned parameter SIZE involved in e.g. *groß - klein* , see Lang 1989). Note that things like MAX, SUB etc. are not mere labels, but theoretical constructs having a clear-cut interpretation within the scope of the DA model outlined in the Introduction. Slightly simplified, the conditions encoded in each of these DAPs may be spelled out like this:

MAX identifies the most extended disintegrated axis of some object **x**, which in turn presupposes there to be exactly one such axis of **x** available (that is what makes *long - short* inapplicable to circles, squares, balls etc.).

SUB identifies either a non-maximal third axis (cf. *thick board, thin slice of bread*) or a non-maximal integrated axis, e.g the diameter of a circular cross section (cf. *thick pole*).

DIST identifies an object axis perceived as inside diameter of a hollow body. Though SUB and DIST identify the same type of axis in terms of geometry, they draw on mutually exclusive perceptual properties viz. permitting or preventing being looked through. Thus, SUB refers to axes determined by solid (parts of) objects, DIST to axes determined by hollow ones.

VERT identifies, if assigned via *high* or *tall* etc. to some spatial object **x**, exactly that disintegrat ed axis of **x** which coincides with the Vertical of PPS.

OBS identifies, if assigned e.g. via *deep* to some spatial object **x**, any disintegrated (non-minimal) axis of **x** which coincides with the Observer axis of PPS.

ACROSS designates a disintegrated object axis which is left unspecified by any of the other DAPs referring to maximality, substance, verticality, or alignment to the Observer axis.

Notice that ACROSS is a stop-gap with respect to both IPS and PPS. Within PPS, ACROSS covers horizontality in that it is assigned to an axis to which neither VERT nor OBS apply; within IPS, ACROSS supplements the parameters MAX and SUB in that it is assigned to an axis to which neither of these applies. In other words, ACROSS represents the overlap of the two categorization grids PPS and IPS, and due to this it provides a major source of ambiguity within and typological variation between languages - as will become clear in sections 3 and 4.

To sum up: it is this small set of semantic parameters DAP = {MAX, VERT, OBS, ACROSS, DIST, SUB} which controls the way in which natural languages assign spatial dimensions and positions to objects[3]. DAP is the stuff which the lexical meanings of dimension terms are made of. Taken as categorized semantic components, the elements of DAP are packaged into more complex representations that e.g. for dimensional adjectives meet the following general format (for details see Lang 1989, Bierwisch & Lang 1989):

(8) $\lambda c\ \lambda x\ [\text{QANT}\ [\text{DIM}\ x]\ =\ [v\ \pm\ c]]$

DA crucially involves gradation and comparison, thus QUANT is a semantic component for a scaling operation which assigns a scale value composed of **v** and **c** to some spatial object **x** with regard to a dimension **d**. The latter is represented by the placeholder DIM, a variable that is to be replaced by the elements of DAP. Though (8) offers a lot of aspects that invite further typological considerations, this paper will focus on the dimension component DIM in discussing the conditions on which the elements of DAP are lexicalized and arranged to lexical fields. Next step in doing this is to take a look at the internal structure of the set DAP = {MAX, VERT, OBS, ACROSS, DIST, SUB} w.r.t. the division of labor between, and the mutual compatibility of, the elements it contains.

---

[3] Note that the approach adopted by Herskovits (1986) and other Cognitive Semanticists is, roughly speaking, observer-centered and situation-based. The approach I am advocating is object-centered and axes-based. This difference in view, I hope, will stimulate the discussion.

1.4. Compatibility conditions.

By claiming that DA to spatial objects is the joint outgrowth of the two categorization grids IPS and PPS, the approach adopted here suitably accounts for the fact that there are various cases in which a given object axis is not identified by a single parameter but by a combination of parameters from both grids. Such combinations of DA parameters occur on the Semantic Level as well as on the Conceptual Level. Here are a few examples for each:

Semantically, the English adjective *tall* comprises a combination of MAX and VERT, though not as a (symmetric) conjunction. The fact that the antonym of *tall* is *short* (not *low*) suggests that MAX and VERT are combined in such a way that the axis referred to by *tall* is identified as the object's maximal axis which is furthermore specified as being aligned to the Vertical.

Conceptually, the Object Schemata for "tree" or "tower" contain as primary (i.e. as defining) entry the complex *max-vert* , which indicates (i) that the objects at issue have a canonical orientation regarding verticality, (ii) that this canonical verticality is bound to the objects' maximal axis.

Besides occurring as conceptually fixed, combinations of parameters can also result from contextual specification. So the primary OS for "pole" contains the entry *max* which suffices to assign things like *the pole is 3 m long* a regular interpretation. The interpretation of *the pole is 3m tall / high*, however, provides the OS for "pole" with the contextually induced verticality specification which also results in a complex entry *max-vert.* This is how a gestalt property (*max*) is turned into a position property (*max-vert*). The fact that specification the other way round (that is, turning position into gestalt properties) is excluded adds further evidence to the analysis of *tall* sketched above and, in general, shows that the relation between IPS and PPS is an asymmetric one.

Now, the claim is this: both the full range of possible objects in OS and the scope of admissible dimensional designations and positional variations of spatial objects (irrespective of being primary or contextually induced) are determined by a small set of compatibility conditions that specify which axial properties may combine. The details are given in Lang 1989, here I will only list the results. Given the interpretation of the elements of DAP in 1.3, we are left with 14 in three groups :

(9) single parameters:
MAX, VERT, OBS, ACROSS, DIST, SUB

(10) admissible combinations (based on compatible axial properties):
MAX-VERT, ACROSS-MAX, MAX-OBS, VERT-OBS ($\measuredangle$ 180°)

(11) inadmissible combinations (due to incompatible axial properties):
*MAX-SUB, *DIST-SUB, *OBS-SUB, *OBS-VERT ($\measuredangle$ 0°)

Note that a combination of parameters - as illustrated with *tall* above - is not a mere conjunction but a more structured complex made up of a basic parameter (left part) and a specificatory one (right part). The combination VERT-OBS in (10) is reserved for concave objects that are canon-

ically aligned to the Vertical and that are specified for a canonic ("river") or a contextually induced perspectivization ("pot") that refers to the same axis but in the opposite direction. The combinations *MAX-SUB,*DIST-SUB,*OBS-SUB in (11) are excluded due to perceptually incompatible axial properties - cf. 1.3 above. The combination OBS-VERT, where the Vertical and the Observer Axis run in the same direction (at 0°), is perceptually quite conceivable but, interestingly enough, does not constitute a conceptually relevant parameter. [4]

To sum up: the claims concerning admissible single and combined DA parameters in (9) - (11) sort out 10 out of 14 and thus lay the ground on which we might now look for universals.

2. Some semantic and lexical universals in the realm of DA

2.1 What are semantic universals supposed to be?

Let me start with a few assumptions on what specific theoretic contributions we should expect from semantic universals in comparison with universals in phonology, morphology, or syntax. Given the two-level approach to meaning worked out in the literature quoted above (Bierwisch, Lang (eds.)(1989) and much subsequent work), I take the status of semantic universals to be that given in (12) and their role in linguistic theory to be determined by the requirements posited in (13) - (16):

(12) Semantic universals are statements on how semantic primes and combinations of primes are encoded into lexically categorized, morpho-syntactically specifiable, hence compositionally suitable building blocks out of which phrasal and sentential structures are formed.

Put in terms of a modular view on grammar, (12) might also be reformulated to the effect that semantic universals are statements on the interface between lexicon-based grammatical structures and conceptual structures (knowledge stored in memory). The attribute ' lexicon-based' is due to the view that meaning in language is necessarily linked with lexical items, that is, with those linguistic units that in a specific way integrate phonological, morphological, syntactic, and semantic information and thus - by projecting this information onto combinatorial structure - constitute the basic elements for grammatical structure formation. In this paper, we will focus on the underlined part of (12) and examine it w.r.t. some requirements that are imposed on semantic universals.

(13) Semantic universals are expected to contribute to clarifying the difference between, as well as the interaction of, linguistically encoded meaning and extralinguistic context information;

(14) Semantic universals are to serve as basis on which (a) notions like ambiguity, polysemy, and unspecifiedness can be distinguished, (b) the various types of inferences can be explained;

---

[4] There are data which prove that this kind of doubly determined axis identification is not utilized semantically. As piece of evidence let me quote the following documented example from a TV report on Cape Canaveral:
(i) *The rocket rose into height and disappeared in the depth of space*
The fact that the visible path of the rocket covers one continuous segment (simultaneously determined by the Vertical and the Observer axis running equidirectionally) cannot be semantically accommodated in one dimension term. Instead, the semantic structure of dimension terms necessitates a way of designating the path of the rocket which construes the relevant projections as concatenated, the point of linkage being marked by a shift in the reference system (from PPS to the orbit).

(15) Semantic universals that state which (combination of) primes is lexicalized, that is, packaged into categorized lexical items and this way put into grammatical structure formation, should draw on independent explanations as far as possible.

With these preliminaries and the definitions of the elements of DAP in mind, we may now move on to formulating some tentative universals under the follwing heading:

2.2 What of DAP is lexicalized ?

Given the decisive role of the compatibility conditions discussed in 1.4, especially the assumptions posited in (9) - (11), there is a universal constraint which suggests itself as it follows from (11). We can put it in two equivalent versions:

(16) Only admissible combinations of elements from DAP are lexicalized (=U-1)

There are no lexical items covering simultaneous reference to axes that would be identifiable by MAX-SUB or DIST-SUB or OBS-SUB or OBS-VERT. (=U-1')

In view of the fact that the supposed combinations MAX-SUB, DIST-SUB, and OBS-SUB would embody perceptually incompatible cues (regarding OBS-VERT cf. FN.3), (16) is not a surprising result. It nevertheless is worth being mentioned as a it illustrates what the requirement to look for independent explanations as posited in (15) is assumed to mean. Furthermore, (16) has some direct implications for lexicalization. It predicts that parameters that draw on mutually incompatible axial properties are lexicalized separately, that is:

(17) If MAX, SUB, DIST, OBS are included in a lexical field of DA, the terms drawing on any subset of them are lexically distinct.

The next two tentative universals draw on the prominence of the axes determined by IPS and PPS and are formulated a positive hypotheses. The first one again in two versions:

(18) The most prominent axes of both IPS and PPS are lexicalized separately (=U-2)

In a lexical field of DA, there are at least distinct items for MAX and SUB as well as for VERT and OBS (=U-2')

Unlike (17), which is a corollary of (16), the statement in (18) is an independent one. It claims separate lexicalization for MAX (from IPS) and VERT and OBS (from PPS) and ACROSS (possibly from either) despite the fact that the axial properties to which these parameters refer to are pairwise compatible - as shown in the admissible combinations in (10). (18) is a plausible tribute to saliency as a source for lexicalization which, of course, has to be proved by massive empirical evidence yet. It is a claim about minimal distinctness in the lexical field of DA, and hence is not at variance with e.g. English having items covering VERT (*high - low*), MAX (*long - short*), and MAX-VERT (*tall*) separately, nor with the possibility that there may be languages that encode OBS (at $\measuredangle$ 90°to VERT) and OBS-VERT (= OBS at $\measuredangle$ 180°to VERT) into distinct lexical items.

The following is a tentative universal concerning the relation between conceptual salience and lexical primes (a lexical prime being a lexical item that is neither a compound nor a derivative):

(19) DA terms drawing on MAX, SUB, VERT, OBS are rendered by lexical primes. (= U-3)

Well, in all of the 15 or so languages I have examined so far (besides the major Germanic, Romance, and Slavonic languages my data base includes Turkish, Korean, Mandarin and Khmer), the field of DA belongs to the core lexicon, so the claim in (19) may be weak. It is meant as an attempt to correlate basic perceptual contrasts with preferences for lexical primeness. I hasten to add for all of (16) - (19) that, though I did not find a counterexample to any of them, all three need further confirmation by massive empirical data. So much for tentative universals of what is (minimally) lexicalized within the field of DA. The scope thus defined leaves room for variations in the internal structure of the lexical field of DA. This is what we will take up now.

3. Typological variation in the structure of the lexical field of DA: the basics

3.1 What are the basics ?

Taking stock of the inventory of single and admissible combined DAP with the help of Fig. 3. below, we observe that - disregarding polar antonyms - there are at least 10 candidates for being lexicalized as basic dimension terms. This number of potential DA terms is much larger than the number of basic dimension terms we observe in most known languages, which ranges between 5 to 8 (again disregarding antonyms). Hence we face the problem of what lexical items cover what subsets of the DAP inventory shown below.

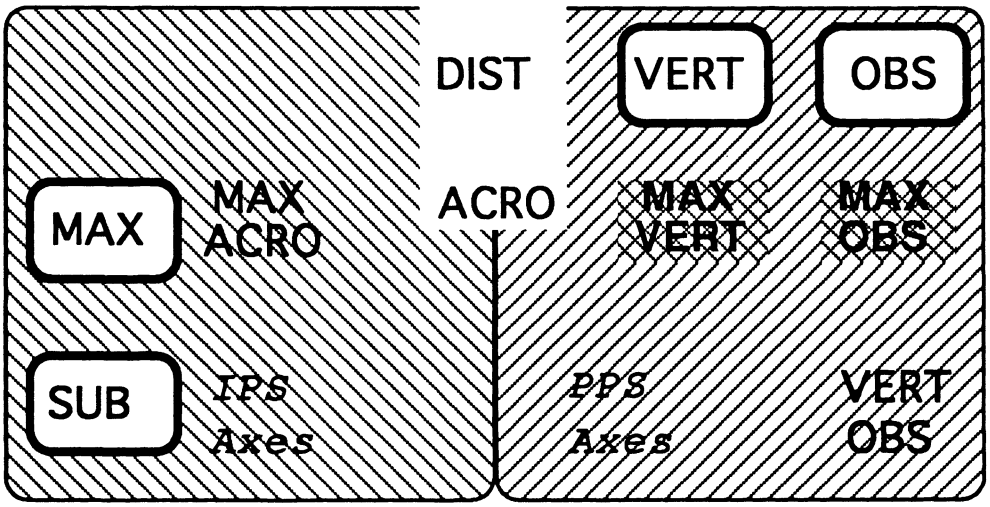


Fig. 3. IPS & PPS Sources of DAP.

As a consequence of this inevitable few-many mapping from lexical items to DAP, we should expect various possible partitions of the set of DAP as regards the scope of lexical coverage. Languages actually vary as to these partitions, and going in search of recurrent patterns and principles explaining them we enter the field of typology.

Taking for granted that, following (19), the two salient DAP of each grid (i.e. MAX, SUB, VERT, OBS - bold encircled in Fig. 3) are rendered by lexical primes, the scope of varying partitions is supposedly confined to the rest. If so, what determines the boundaries of a given partition?

Presumably, it is not the boundary separating IPS and PPS as we observe (indicated by the hatching) two areas which blur such a simple division:

(a) the combinations MAX-VERT and MAX-OBS embody parameters from both IPS and PPS. As there are lexical primes that cover such a combination (recall *tall*), the boundary between IPS and PPS hence does not necessarily serve as a demarcation line in the lexical field;

(b) the parameters ACROSS and DIST are somehow indeterminate between IPS and PPS (cf. 1.2 and 1.3 above), which does not make them clear-cut boundary posts either.

Given that dimension terms rest on conditons for identifying object axes, a cross-linguistic examination of the lexical field of DA suggests that the assignment of dimensions to objects inter alia follows two different **strategies** (20), (21) and an invariant principle (22) which in turn seem to determine the partition of the lexical field of DA .

(20) proportion-based strategy: the relative size of an object's axial extensions **a**, **b**, say **a** > **b**, determines the assignments of dimensions to objects.

(21) observer-based strategy: the condition whether or not **a** = OBS, **b** = ACROSS to OBS (or vice versa) determines the assignment of dimensions to objects.

Independent of (21) and (22), the fact that the Vertical is the domint axis of our spatial orientation (cf. (5) in 1.1) lays the ground for another invariantly observable principle which reads:

(22) the Vertical prevails: if a significant (i.e. non-minimal) object axis **a** coincides with the Vertical of PPS, then this coincidence will determine what dimension **a** is assigned.

To illustrate (20) - (22), imagine a writing desk in normal position sized **a** = 2m, **b** = 1m, **c** = 0.80m. Whatever the proportion of its axial extensions, (22) will pick out its canonic top-bottom extension, say **c,** and reserve it for being identified by a term covering VERT, e.g. *It is 0.80 m high. / in height.* Now, if the remaining extensions **a**, **b** are described by *It is 2m long / in length and 1 m wide / in width,* (20) applies. If the same extensions are described by *It is 2m wide / in width and 1 m deep / in depth,* (21) applies. Actually, English can make use of both strategies (which - as we will see below - is a typologically relevant feature) though not simultaneously in one and the same construction: **It is 2m wide / in width and 1 m wide / in width* is clearly out. The unacceptability of such expressions leads us to a constraint which - being a corrollary of some more general conditions on wellformed coordinate structures - can be narrowed down regarding DA to this:

(23) Uniqueness constraint: in an instance of identifying distinct axial extensions **a, b, c** of some object x, a dimension term t may apply only once.

With (20) - (23) in mind we are prepared to approach the typology of lexicalization patterns.

3.2 How languages partition the lexical field of DA - a typology

Taking (20) and (21) as key factors in determining possible partitions of the lexical field of DA and examining DA data from a sample of 15 languages (a selection of which will be exemplified below) allows us to reach the following conclusions. Typologically, languages differ as to whether or not that subset of the lexical field of DA which covers ACROSS

(24) (a) is clearly determined by the proportion-based strategy (e.g. Mandarin Chinese) or
(b) is clearly determined by the observer-based strategy (e.g. Korean) or
(c) is determined by a conflation of both strategies (German, English, French, Russian).

Note that (24)(a-b) allows for both strategies to apply for disjoint subsets of the same lexical field, the crucial point is whether or not both are conflated on the same lexical items. Moreover, the languages in (24)(c) can be ordered or scaled as to the relative share of either strategy they involve. Thus - as will be shown immediately - German is number 1 on the scale of observer-basedness, Russian is number 1 on the scale of proportion-basedness, English and French are in between.

In the following, I will illustrate (24) by data gathered with the help of two sets of elicitation tests with native informants. I will start with the conflated type (24)(c) as it allows to introduce the experimental design on the basis of more familiar languages and then move on to show in which way each of the other types (24)(a) and (24)(b) significantly differs from English or German.

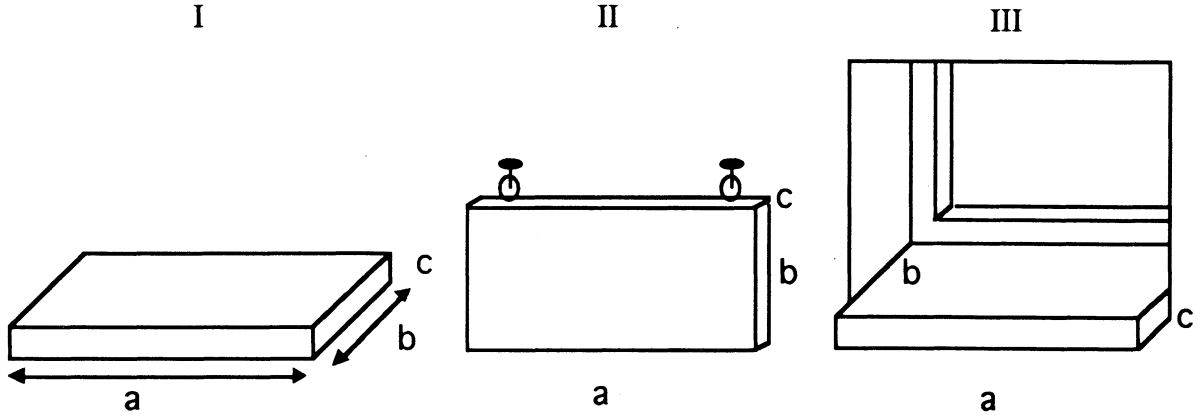
Note that the search for typologically relevant lexicalization patterns is made within the scope of possible variations that is delimited by the (tentative) universals (16) - 19) stated in 2.2. Hence, any typological assumption we are going to formulate is likewise subject to the requirements (13) - (15) posited in 2.1. So, as regards the lexical coverage of DA terms, the elicitation tests should be arranged in such a way that their foreseeable results are relevant (a) to clarifying the relation between encoded meaning and context information, (b) to deciding on ambiguity, polysemy etc., (c) to accounting for various types of inferences. I will comment on these aspects while presenting the tests. The elicitation tests reported on this paper are tasks of naming object extensions.[5]

3.3 ACROSSing the board in English and German

At first I will report on a tried and tested naming task the results of which have proved to be especially telling as regards the methodological requirements repeated above. Subjects were presented a picture showing a board with constant size in three different spatial settings (I - III) -

[5] The data bases underlying the analyses presented in Lang 1989, Lang 1990a,b were obtained with the help of a larger set of elicitation tests which include tasks like object guessing (*x is wide, deep, and high. What might x be* ?) and acting out tasks where subjects were given an object of fixed size, say, a book or a brick, and were asked to position the object according to its possible description by sentences like *The brick is 24cm long, 11cm high, and 7cm wide* . These tests are not discussed in the present paper as they were made only for German and English so far. It goes without saying that they deserve to be carried out for other languages — provided the approach advocated here should prove to be fruitful.

see. Fig. 4. below. Subjects were asked to name the object's axial extensions, which were labeled by the letters **a**, **b**, **c**. Subjects were given hints on plausible measures of the object, e.g. **a** = 1m, **b** = 30cm, **c** = 3cm. For German and English, which in this case do not differ, the results are those listed in (25) and (26) under the respective settings shown in **Fig. 4.**



	I	II	III
(25)	a = lang b = breit c = dick	a = breit b = hoch c = dick	a = breit b = tief c = dick
(26)	a = long b = wide c = thick	a = wide b = high c = thick	a = wide b = deep c = thick

Well, that extension **c** is constantly labeled *dick / thick* is not surprising in view of the claims about SUB (which is encoded by *dick / thick*) made in section 1 above: (9) states that SUB does not enter combinations, (17) claims that SUB is lexicalized distinctly from MAX, OBS, DIST, (19) claims SUB to be packaged into terms that occur as lexical primes.

With extensions **a** and **b**, however, we observe correlative changes of DA terms which obviously (i) centre on how the stop-gap parameter ACROSS is lexicalized (or, put the other way round, what is semantically covered by the terms *breit / wide*); (ii) depend on the spatial settings the board is in. Now, what do (i) and (ii) tell us about the lexical coverage of the terms *breit / wide* ?

As for (i), there are two points to be made. The first point is that - given the 10 admissible (combinations of) parameters in (9)-(10) and in Fig. 3 - *breit /wide* cover the single parameter ACROSS (25)(I) but also the combination MAX-ACROSS (25)(II - III). Looking for the conditions on which these assignments take place leads us immediately to the second point.

Recall that the parameter ACROSS serves as a stop-gap in PPS and/or IPS as it lacks independent perceptual support, which in turn implies that ACROSS comes into play only as dependent on other parameters which do draw on independent perceptual support. This dependency is what defines the respective place of the terms that cover ACROSS in the lexical field of DA. We will see that the typology of languages outlined in (24)(a - c) essentially rests on what strategy takes care of ACROSS and in what way, that is, the proportion-based or the observer-based in disjoint subsets of the lexical field or both conflated onto one subset, which leads us back to German and English.

Now, given that there is no independent defining spatial property according to which *breit / wide* are assigned to physical objects, an object extension **d** to which *breit / wide* are assigned is determined in relation to some other object extension **d'**, where **d'** is identifiable independently. Let's try then to interpret the data in (25/26) against the background of the claims made so far.

In setting I, the assignments **a** = *lang /long*, **b** = *breit /wide* follow the proportion-based strategy (20), that is, ACROSS is determined in relation to some other axis **d'** which is identified as the maximal (i.e. **d'** = **a** = MAX) by *lang / long*. The occurrence of the latter is, as (19) predicts, reliably indicative of maximality being the defining property of the given extension.

In setting II, the assignment **b** = *hoch /high* follows from (22) - "the Vertical prevails" - and as such yields an independently identified axis onto which ACROSS can be hooked (i.e. **d'** = **b** = VERT), thus **a** = *breit / wide.* But what about the underlying strategy? If **a** in setting II were assigned proportion-based, it would be labeled *lang / long*, which is, however, unacceptable in both languages if paired with **b** = *hoch / high* . Hence, the assignment **a** = *breit /wide* in setting II (where it is paired with **b** = *hoch / high*) is due to the observer-based strategy as well. This claim is supported by the fact that normally the axes of PPS are orthogonal to each other (OBS at $\measuredangle$ 90° to VERT and ACROSS at $\measuredangle$ 90° to OBS) as has been claimed in (6) and (7) above.

In setting III, the assignments **a** = *breit / wide*, **b** = *tief / deep* follow the observer-based strategy (21), that is, ACROSS is determined in relation to some other axis **d'** which is identified as being aligned to the Observer axis (i.e. **d'** = **b** = OBS) by *tief / deep*. The occurrence of the *tief / deep* is, as follows again from (19), a reliable indicator of observer-basedness.

Concerning the aspects of ACROSS noted in (i) above, we are now ready to state the relevant partition of the lexical field of DA in German and English. The complete patterns of DA terms, which also show differences between the two languages beyond (27), will be given in 4.3 below.

(27) If *breit /wide* do not cover MAX but are assigned relative to **d'** = MAX, they are assigned proportion-based, otherwise they are assigned oberver-based.

Next, in order to take up (ii), let's take a look at the part played by the spatial settings in (25/26). In (I) the board is conceived as a freely movable object, hence the dimensional terms assigned to it refer to its inherent gestalt properties which allows the proportion-based strategy to apply. Note that due to referring to the object as such, the assignments made in (I) will also hold for settings (II) and (III), though not vice versa. This fact is important for the inferences to be accounted for. In (II) the board can be conceived as having undergone an orientation towards the Vertical due to which it can be assigned position properties (to be hanging, to have a height etc.). This is what the dimension terms **b** = *hoch/high*, **a** = *breit / wide* in (II) refer to.

Finally, in (III) the board can be conceived as part of the window niche such that it is assigned position properties it inherits from the surrounding macro-object, i.e the depth and width of the niche are transferred to the board thus specifying its inherent properties. This is what explains the assignments **b** = *tief /deep* , **a** = *breit / wide* in setting (III).

Having discussed the test so far, it is time for a few remarks assessing its heuristic value. Methodologically, there are at least three aspects that make this test a powerful tool in DA research.

First, it shows a board of constant size in three settings that reflect the board's increasing integration into the spatial environment. In fact, the board as shown in Fig. 4 can be claimed to be distinctly conceptualized in the three settings[6], which, in turn, will reveal to what extent the DA terms used in each of the settings are context-dependent or, put the other way round, what the DA terms used in (II) and (III) induce as contextual specification.

Based on this, the test yields an interesting means to assess the equivalence of situational and linguistic contextual information. The conditions governing the assignment of *breit / wide* that were illustrated in Fig.4 (I - III) by means of the non-verbal contextual settings can just as well be illustrated by means of sentences like (28)(I - III) which, drawing on the Uniqueness constraint in (23), provide exactly the same contextual information on **d'** that is needed to assign *breit / wide* .

(28) Das Brett ist lang und breit genug, aber zu dünn — [**d'** = MAX, **b** = ACROSS
The board is long and wide enough but too thin — as in (25/26) (I)]

Das Brett ist breit und hoch genug, aber zu dünn — [**d'** = VERT, **a** = MAX-ACROSS
The board is wide and high enough but too thin — as in (25/26) (II)]

Das Brett ist breit und tief genug, aber zu dünn — [**d'** = OBS, **a** = MAX-ACROSS
The board is wide and deep enough but too thin — as in (25/26) (III)]

We thus get an instrument that allows to account for the impact of non-verbal context information to the extent that it can be captured by the mutual determination conjoined DA terms impose upon each other within a sentence. (By the way, this equivalence of situational and linguistic contextual information can only be accounted for by representations on the Conceptual Level (cf. Fig. 1), which provides a strong argument in favour of the two-level approach mentioned in section 1.)

Second, the test provides us with a clear-cut example by means of which we can check the inferential relations holding between gestalt and position properties of spatial objects. Note the valid inferences in (29). The sentences contain measure phrases in order to secure the constancy of the object extensions at issue.

(29) (a) The board is 1 m wide, 30 cm deep → The board is 1 m long, 30 cm wide
(b) The board is 1 m wide, 30 cm high → The board is 1 m long, 30 cm wide
(c) The board is 1 m long, 30 cm wide ↛ The board is 1 m wide, 30 cm high / deep

This proves that an object's inherent gestalt properties can be validly inferred from its contextually induced position properties but - as witnessed by (29)(c) - not the other way round. The general pattern underlying these inferences is that of de-specification (see Lang, Carstensen, Simmons 1991).

---

6 The fact that in many languages the board in (III) has a special name pointing to its spatial integratedness might be taken as an additional hint in that direction. So unlike the German compound *Fensterbrett* , which keeps *Brett* as head, the English compounds *window-sill, windowledge* have heads that are spatial meronyms; Russian *podokonnik* is derived from *pod* [under] *okno* [window].

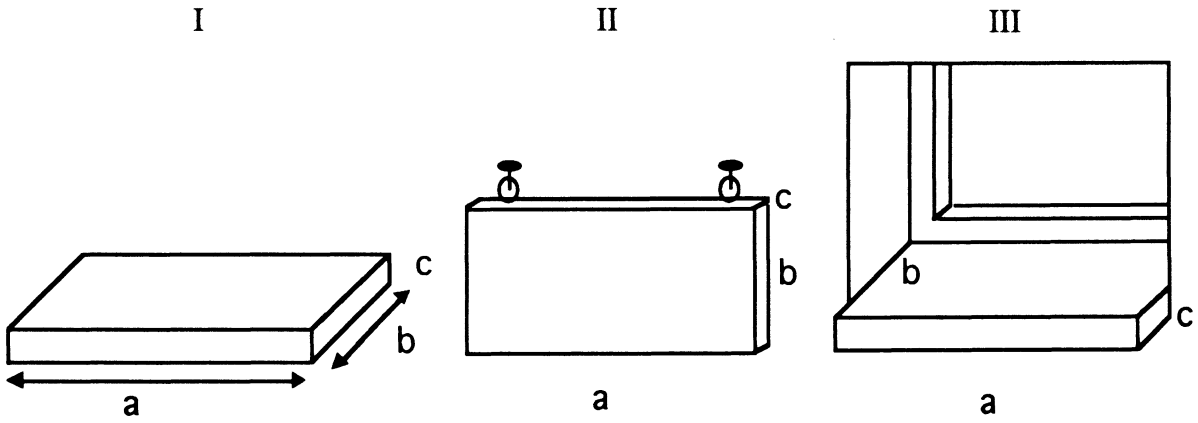
Third, if linked with the typological assumptions in (24), the test provides us with a useful diagnostic for ambiguity. Recall that German and English are of the mixed type regarding the division of labour between proportion-based and observer-based assignments as spelled out for *breit / wide* in (27). Given this we will not only expect, and empirically confirm, that isolated from the context a sentence like (30) is ambiguous as to what extension *breit / wide* refer to,

(30) Das Brett ist 50 cm breit / The board is 0.5 m wide / in width

we can now also exactly determine the range and the source of the different assignments the terms *breit / wide* are able to cover. Moreover, if the ambiguity of *breit / wide* in (30) is an outcome of the fact that German and English are of the mixed type, we would expect the translation of (30) into a language belonging to the "pure" type, say Chinese or Korean, should turn out to be non-ambiguous in this respect. In fact they are - as we shall see in the sequel.

3.4 ACROSSing the board in Chinese, Korean, and Russian

Presenting the naming task in Fig. 3 to speakers of Mandarin Chinese reveals that this language - as regards the lexical encoding of ACROSS - is exclusively proportion-based.[7]



(31)

I	II	III
a = cháng	a = cháng/ *kuān	a = cháng/ *kuān
b = kuān	b = gāo	b = kuān / *shēn
c = hòu	c = hòu	c = hòu

As distinct from (25/26), we observe here that extension **a** is constantly labeled by the term *cháng* that encodes exlusively MAX (as do *long / lang*), while **b** in setting I and III is reserved for *kuān* which is strictly confined to cover ACROSS in relation to **d'** = MAX, and hence is unavailable for **a** in (31) on principle (indicated by * *kuān*).

In setting II, the assignment **b** = *gāo* ['high'] follows from (22) - "the Vertical prevails", which does not interfere with the proportion-based assignment **a** = *cháng/ * kuān* in setting II. If the hanging board is revolved by 90°, the assignments are **a** = *gāo* ['high'], **b** = *kuān / * cháng*, which again confirms (22) as an independent principle.

---

[7] The data shown in (31) confirm observations on Chinese extension terms by Zubin & Choi (1984) and Li (1988), two stimulating papers from which I profited much, not the least because they adhere to a different framework.

The fact that **b** = *shēn* [' deep'] is excluded in any of the settings yields another piece of evidence that Mandarin does not have an observer-based strategy at its disposal. Of course, as predicted by (18) and (19), Mandarin does have a primary term for depth viz. *shēn* ['deep']. However, *shēn* does not at all interact with the conditions of encoding ACROSS but is selectionally restricted to concave or hollow objects. This selectional restriction on *shēn* is again indicative of proportion-basedness (see section 4.1 below.)

Finally, that extension **c** is constantly labeled by one and the same term is not surprising, we would expect this from the comments made on SUB with (25/26). But there is an interesting addition to be noted: unlike German or English, Mandarin Chinese makes subtle distinctions in the lexical items that cover the parameter SUB, so *hòu* is selectionally restricted to identifying a non-maximal disintegrated third axis of an object (i.e. to the first part of the definition of SUB given in 1.3). The issue of lexical granularity of DA terms will be taken up again in section 4 below.

In sum, all observations we can squeeze out of the data in (31) testify to the proportion-based strategy we assumed to be a typological feature of Mandarin. Accordingly, the non-ambiguity prediction regarding *kuān* is borne out: while *breit /wide* in (30) are ambiguous, *kuān* in (32) is not.

(32) Zhè kuài mùbǎn kuān wǔ-shí límǐ
this CL board wide 50 cm
This board is 50 cm wide/in width [i.e. in its secondary extension !]

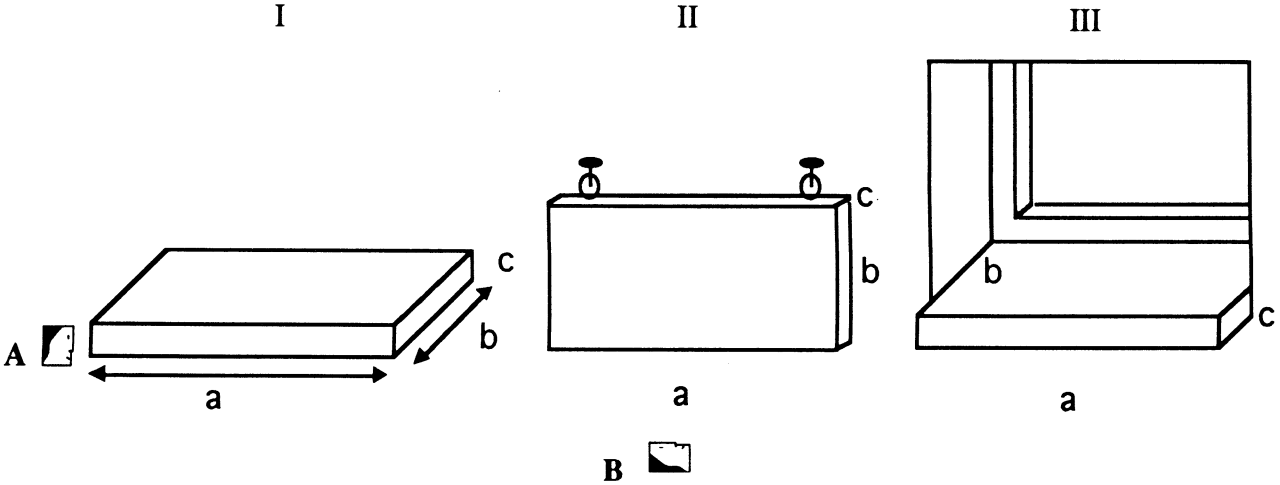
Likewise, the inferences (31) allows for are entirely different from the ones (25/26) do. In order to reveal the semantic consequences of proportion-basedness, I will repeat the English examples of (29) and contrast them with their non-existent equivalents in Mandarin (for brevity' s sake I insert European style measure phrases):

(29) (a) The board is 1 m wide, 30 cm deep → The board is 1 m long, 30 cm wide
(b) The board is 1 m wide, 30 cm high → The board is 1 m long, 30 cm wide
(c) The board is 1 m long, 30 cm wide ↛ The board is 1 m wide, 30 cm high / deep

(33) (a) Zhè kuài mùbǎn kuān 1m, *shēn 30 cm ↛ Zhè kuài mùbǎn cháng 1m, kuān 30 cm
(b) Zhè kuài mùbǎn *kuān 1m, gāo 30 cm ↛ Zhè kuài mùbǎn cháng 1m, kuān 30 cm
(c) Zhè kuài mùbǎn cháng 1m, gāo 30 cm → Zhè kuài mùbǎn cháng 1m, kuān 30 cm

As a matter of fact, the inference patterns in (29) and (33) turn out to be nearly complementary. First note that strictly proportion-based Mandarin cannot provide literal equivalents of the antecedent sentences in (29)(a, b). So it simply lacks the premises needed for this type of inference. Recall that this type of inference draws on the de-specification of oberver-based specifications — and this is what Mandarin does not provide. Second, while Mandarin does not have an analogue of (29)(c) either, it does allow for an inference like (33)(c), which draws on the de-specification of the verticality feature contextually induced according to (22).

Let's now turn to Korean, which works in the opposite way — at least to the extent that oberver-basedness is concerned. In fact, the lexical field of DA terms in Korean (which I quote here in their basic nominal form) also contains a subset of terms (*kili* ['long'], *phok* ['wide']) that resemble Chinese *cháng* , *kuān* in being proportion-based items for an object's maximal and secondary (dis-integrated) axis, respectively.[8] But Korean DA terms contain another subset which is clearly observer-based in that, following Zubin & Choi (1984: 337), the "spatial terms *kalo* and *selo* [...] pick out the edges of a surface which are across and in line with the observer's visual field, respect-ively, with no regard for the relative extension of theses edges." Rephrased in our terminology: these terms are correlatively and context-dependently assigned to non-minimal axes in such a way that *kalo* is reserved for covering ACROSS in relation to **d'** = OBS, *selo* for covering OBS in relation to **d"** = ACROSS. Hence, in order to put the test in Fig. 3 to work, we have to add hints on the respective positions taken by the observer - indicated by the faces and marked by A or B. The proportion-based versions are listed under C.



(34)		I	II	III
	(A)	a = selo b = kalo c = khuki	a = selo b = nophi/kalo c = khuki	a = selo b = kalo c = khuki
	(B)	a = kalo b = selo c = khuki	a = kalo b = selo/nophi c = khuki	a = kalo b = selo c = khuki
	(C)	a = kili b = phok c = khuki	a = kili b = nophi c = khuki	a = kili b = phok/kiphi c = khuki

Extension **c**, as we would expect, is constantly labeled by the term *khuki* ['thick'] covering SUB as defined in 1.3 and does not interfere with taking proportion-based or observer-based strategy. But what is worth noting is the complementary distribution of *kalo* - *selo* onto the extensions **a** and **b** within and between the settings (A) and (B). The additional option for **b** = *nophi* ['high'] in II for all of (A- C) follows independently from (22). The only structural difference of the data in (C) to their likewise proportion-based counterparts in Mandarin is the option for **b** = *kiphi* [' deep'].

---

[8] See also Zubin & Choi (1984). The Korean data in this paper are due to my informant Byong-Rae Ryu (Tübingen), the transcription system for Korean used in the examples below is based on Yale Romanization.

The observation that the two subsets of the lexical field that cover ACROSS are strictly disjoint as concerns source and distribution is furthermore confirmed by the following fact: any conjoined co-occurrence of elements from either subset, that is, any combination of *kalo - selo* and *kili - phok* in a coordinate construction, is ruled out.[9] This homogeneity constraint can be construed as a language-particular tightening of the general Uniqueness Constraint in (23).

Regarding predictable ambiguities, we are faced with two cases as we might have guessed. The proportion-based version of referring to the board's ACROSS axis in (34) is non-ambiguous

(35) Ku-nelphanci-nun phoki il meta ita.
DET.board.TOP width.SUBJ 1 m DECL
"This board is 1 m wide/in width" [i.e. in its secondary extension !]

whereas observer-based versions like (35), if presented without contextual cues for (A) or (B), is ambiguous (or unspecified) as regards reference to extension **a** or **b.**

(36) Ku-nelphanci-nun selo-ka/kalo-ka il meta ita.
DET.board.TOP observer-axis.SUBJ / across axis.SUBJ 1 m DECL
"This board is 1 m wide/in width/long/in length"

Finally, the inferential behaviour of the data in (34) can be easily extrapolated from what has been stated so far: if the absolute measures of the object's extensions are as required, there is room for valid inferences from oberver-based assigned axes to proportion-based ones, that is, from an object's contextually induced position properties to its inherent gestalt properties. Thus, (37) proves to be valid along the lines of de-specification illustrated by (29) in 3.3.

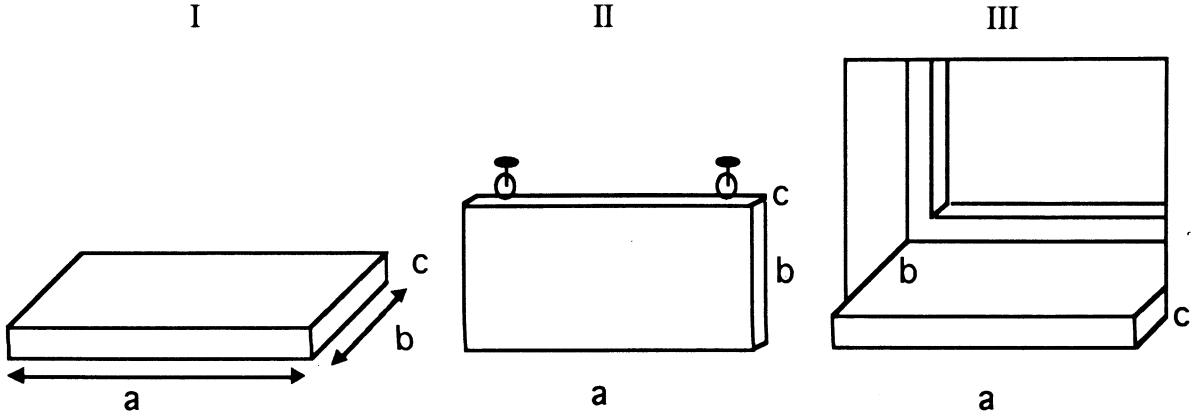
(37) Ku-nelphanci-nun kalo-ka 1 m, selo-ka 0.3 m ita.
→ Ku-nelphanci-nun kili-ka 1 m, phoki 0.3 m ita.
"this board's across axis is 1m, its observer-axis is 0.30 m"
→ "this board is 1m in length, 0.30 in width"

To conclude: as regards the lexical coverage of ACROSS, Mandarin Chinese is confined to a subset of terms that exclusively rest on proportion-based DA, Korean has two disjoint subsets of terms, one of them drawing on proportion-based DA, the other on oberver-based DA, German and English have one subset of terms on which both strategies are conflated. So much for the coarse typology sketched in (24)(a - c).

Let me just add one more example to illustrate the finer-grained sub-typology within the group of "conflating" languages. On the whole, the field of German DA terms shows more features of observer-basedness than of proportion-basedness, in Russian the opposite holds. This can be partially revealed by (38), the Russian terms are given as derived nominals:

---

9 This does not exclude sentences with terms from both subsets occurring in distinct syntactic positions as in:
(i) Ku-nelphanci-nun selo-ka kilta. "The observer-axis of this board is long"
(ii) Ku-nelphanci-nun selo-ka selo-pota kilta. "The observer-axis of this board is longer than its across-axis"



(38)	Russian	a = dlina	a = dlina/shirina	a = dlina/*shirina
		b = shirina	b = vysota	b = shirina/*glubina
		c = tolshchina	c = tolshchina	c = tolshchina
	German			
		a = Länge	a = Breite/?Länge	a = Breite
		b = Breite	b = Höhe	b = Tiefe
		c = Dicke	c = Dicke	c = Dicke

The relevant point is in setting III: while German opts for observer-based assignment, i.e. **b** = *Tiefe* encodes OBS and thus provides a suitable **d'** to which ACROSS in **a** = *Breite* can be hooked on, Russian sticks to the proportion-based assignment **a** = *dlina* ['length'], **b** = *shirina* ['width'], the label **b** = *glubina* ['depth'] is rejected as unacceptable. What does this tell us about the place of Russian on the scale?

First, unlike Mandarin it does not restrict ACROSS to being dependent on **d'** = MAX only, it also allows for a **d'** = VERTas witnessed by the assignments in II. Here **b** = *vysota* ['height'] (which follows independently from (22)) leaves room for either having proportion-based **a** = *dlina* ['length'] or - less preferred - for **a** = *shirina* ['width'].

Second, like in Mandarin but unlike in German and English, in Russian the term encoding OBS, i.e. *glubina*, is selectionally restricted to concave or hollow objects (which a board apparently does not belong to, even if integrated in the window niche in setting III).

So, the range of ACROSS axes to be covered by Russian *shirina* ['width'] is confined to those axes **d** that are hooked on **d'** = MAX or **d'** = VERT whereas German and English have an additional option for **d'** = OBS, the latter being obviously linked to the fact that *Tiefe / depth* are not selectionally restricted to concave or hollow objects but encode OBS without restrictions.

This observation reveals an interesting facet of cross-linguistic variation in DA terms: languages that in accordance with (19) have the same number of lexical primes covering MAX, SUB, VERT, and OBS, respectively, may still differ as to the selectional restrictions imposed on these terms. So, selectional differences should be included in the list of typological parameters.

In the next section we will scrutinize the lexical field of DA terms for further aspects of the typology outlined in (24) while at the same time presenting another useful test setting.

4. Typological variation in the structure of the lexical field of DA: more details

In section 3 we focussed on discussing what strategies take care of covering the parameter ACROSS and on pinpointing the effects of the various options on the structure of the lexical field of DA. Now, there is a lot of evidence that the choice between, or the interaction of, proportion-based and observer-based DA has more effects on determining the lexical coverage of DA terms. Space limitations prevent me from presenting the full particulars that might be adduced to back up this claim. Instead, I will proceed like this: 4.1 presents a list of characteristics of observer-basedness and proportion-basedness, respectively; in 4.2 these criteria will then be illustrated by applying them to data that were obtained from another test setting; 4.3 will summarize the aspects discussed so far by presenting some full lexicalization patterns of DA terms.

4.1 Characteristics of observer-basedness vs. proportion-basedness

The following sets of characteristics were abstracted from a large amount of data. They are listed as informal descriptions of features that are symptomatic of the impact on the lexical field of the observer-based or the proportion-based strategy in DA. To easy later reference, I will number them, L is a variable for languages. To begin with, the impact of the **observer-based strategy** on the internal organization of the field of DA terms in L can be diagnosed from (O-1) - (O-4) below, which - as far as I can see - do not display any intrinsic order.

(39) The term(s) that cover(s) OBS in L determine(s) the use and the range of interpretations of the term(s) that cover(s) ACROSS in L. (= O-1)

(40) The term(s) that cover(s) OBS in L is/are <u>not</u> selectionally restricted to concave or hollow objects. (= O-2)

(41) For a specified class of objects C, L allows for a ' high' - ' deep' alternation in designating some object axis **d**. (= O-3)

(42) In L, verticality assignment can absorb maximality assignment, that is, the parameter-combination MAX-VERT is covered by the terms of L. (= O-4)

In view of the basics discussed above, only (O-3) and (O-4) deserve some comments. The ' high' - ' deep' alternation mentioned in (O-3) occurs when an object's (primary or contextually induced) vertical axis [' height'] is contextually specified as being aligned to the observer-axis. Technically, if in the object's OS an entry *vert* is contextually combined into *vert-obs* thus specifying that the axis **d** referred to as vertical is being looked at in the opposite direction (OBS at $\measuredangle$ 180°to VERT), cf.:

(43) Der Topf ist zu <u>hoch</u>, um ins Regal zu passen, aber nicht <u>tief</u> genug für die Pute
The saucepan is too <u>high</u> to fit into the shelves but not <u>deep</u> enough for the turkey

The delimitation of the class of objects C is dependent on whether or not (O-2) and (O-3) jointly hold, as is the case for German, where the range of the ' high' - ' deep' alternation is rather wide.

(O-4) refers to the alternative left open if (22) ("the Vertical prevails") applies. If the maximal axis of an object happens to be also the vertical (e.g. a corner cupboard of 2m height), L has to decide

(a) whether the term assigned to this axis is reserved for verticality alone, maximality being transferred to the secondary axis and accounted for in a separate assignment or

(b) whether the term assigned to this axis designates verticality while absorbing maximality in such a way that no other axis is available for the term that assigns maximality.

The option (a) is what we observe in Mandarin or Russian, where the cupboard is assigned *vysota* ['height'], and independently *dlina* ['length'] and *shirina* ['width'], whereas option (b) is typical for German and English, yielding the assignments *high* , *wide* (**long*) and *deep* , or even *tall*, *wide*, and *deep,* where in *tall* the absorption of maximality by verticality is explicitely lexicalized.

Next, let's look at the characterics of **proportion-based strategy** being operative. As one might guess, (P-1) will be supplementary, while (P-2), (P-3), (P-4) will be complementary to the symptoms (O-1) - (O-4) and hence self-explaining.

(44) The term(s) that cover(s) MAX in L determine(s) the use and the range of interpretations of the term(s) that cover(s) ACROSS in L. (= P-1)

(45) The term(s) that cover(s) OBS in L is/are selectionally restricted to concave or hollow objects. (= P-2)

(46) If L allows for a ' high' - ' deep' alternation in designating some object axis **d**, then the class of objects C for which it holds is more constrained than in (O-3) (= P-3)

(47) In L, verticality assignment is separated from maximality assignment. (= P-4)

Having the two sets (O-1) - (O-4) and (P-1) - (P-4) at our disposal, we may now take further steps towards a finer-grained specification of typological variation in the lexical field of DA terms. As regards the internal structure of this field, a language L can be scaled by stating which of (O-1) - (O-4) and/or of (P-1) - (P-4) can be proved to apply in L. Moreover, the characteristics expounded above provide us with a suitable means to account for ambiguities and for lexical gaps. This will be shown in the next section where I proudly present another elicitation test.

4.2 The *staircase* case

The experimental design of this naming task is again quite simple. Subjects were presented the picture of a staircase in Fig. 5 below and asked (a) to name the extensions of the first step (shadowed), (b) to name the dimensions of the staircase as a whole. Subjects were hinted to answer twice, while imagining themselves (i) going upstairs, (ii) going downstairs (big arrows).

The test is valuable in several respects: it allows to check the availability in L of the ' high' - 'deep' alternation, it reveals the degree of spatial integration of a step into the macro-object "staircase", and - based on this - it is a diagnostic for inferences that draw on DA inheritance from parts to wholes and vice versa.

4.2.1 Coarse Typology. In order to show how the typology in (24) can be reconstructed in terms of (O-1)-(O-4), (P-1)-(P-4), I will first present data from Mandarin, Russian, German, and Korean. The data are arranged in Fig. 5 in such a way that it reflects the underlying typology.

(48) Mandarin

a = gāo	['high']
b = cháng	['long']
c = kuān	['wide']

(49) Russian

a = vysota	['height']
b = dlina	['length']
c = shirina	['width']

(50) German

a = hoch		a = tief	
b = breit		b = breit	
c = tief		c = ???	

(51) Korean

a = nophi	['height']	a = kiphi	['depth']
b = kalo	['across-axis']	b = kalo	['across-axis']
c = selo	['observer-axis']	c = selo	['observer-axis']

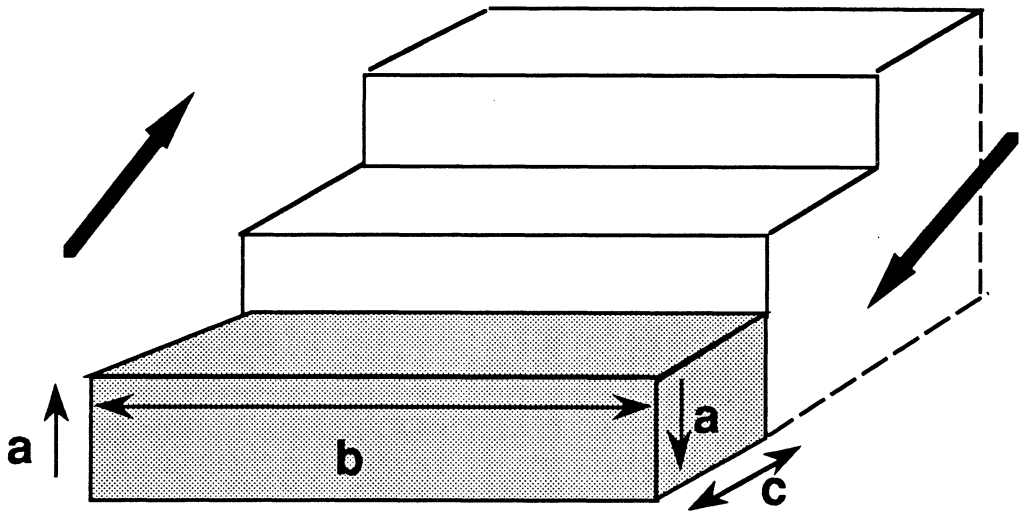


Fig. 5 Naming the extensions of a step of a staircase

Note that (48) marks the only way in which Mandarin can dimensionize a step of a staircase. This provides additional evidence for the claim that Mandarin DA terms are strictly proportion-based. Let's check this by running through the Ps: (P-1) yes, (P-2) yes (a step is not a concave object), (P-3) yes (no 'high' - 'deep' alternation with non-concave objects), (P-4) yes (verticality assignment [*gāo*] is separated from maximality assignment [*cháng*]. So all Ps but no O apply. What is more, in such a clear-cut case of proportion-basedness we would not expect any ambiguities or gaps, and in fact there are none. Russian is similar except that (P-1) is supplemented by (O-1), i.e. the option to assign ACROSS relative to VERT — see (38) II), which may apply in (49), too.

The German assignments in (50) are as expected if assigned going upstairs (left set) but reveal an hitherto undetected gap if assigned going downstairs (right set c = ???). How come? Checking the Os we obtain (O-1) yes, (O-2) yes (both proved by the left set); (O-3) yes (witnessed by **a** = *deep* in the right set), (O-4) yes (not evidenced by Fig. 5 but easy to imagine: if **a** were the maximal the steps would be uncomfortable but the assignments would be the same). So what's wrong?

The gap in (50) is caused by two independent but converging factors. First, as will be recalled, German *tief* covers OBS and OBS-VERT, the latter being the basis for the 'high' - ' deep' alternation we observe in (50), the former causing an interference of (O-1) and (O-3) w.r.t. *tief*. This is the language-particular factor.[10] Second, the Uniqueness Constraint (23) is operative and thus prevents *tief* from occurring twice in (50) despite the fact that it would refer to distinct extensions. This is the universal factor which expectedly prevails.

Finally, (51) shows the observer-based subset of Korean DA terms in two versions. How does this subset score? (O-1) yes (that is the gist of the *selo-kalo* correlation), (O-2) + (O-3) yes (step is in class C), (O-4) no (verticality assignment is separate from maximality assignment). Why are there no ambiguities or gaps? Well, unlike in German, in Korean (O-3) does not interfere with (O-1) as the 'high' - ' deep' alternation [*nophi - kiphi*] is independent of the assignment of *selo-kalo* . So no ambiguity arises nor is there room for a gap.

4.2.2 Finer-grained typology. Another set of staircase data is that in (52) - (55). The point here is the variations we find among languages that are near cognates in i.a. using terms which etymologically share the same common Slavonic origin. Russian, as shown in (48), is high on the scale of proportion-basedness, whereas Polish and Slovak, while using the same adjectival terms, display a tendency towards observer-based assignments — presumably under the influence of German.

(52) Russian

a = vysokij	['high']		
b = dlinnyj	['long']		
c = shirokij	['wide']	/*glubokij	['deep']

(53) German

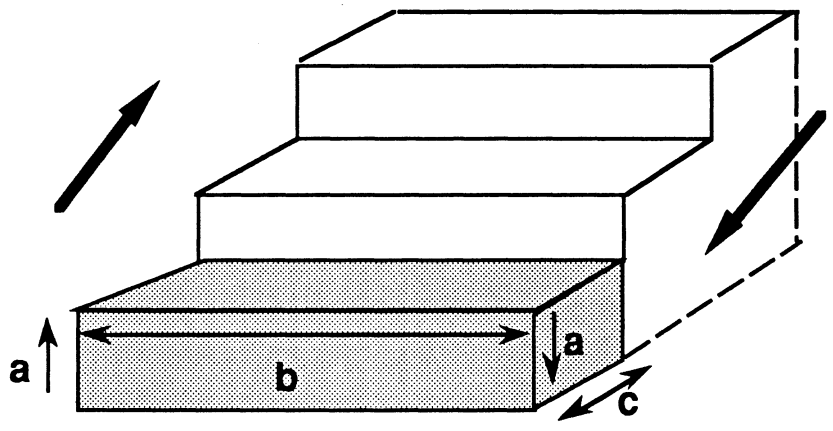
a = hoch	['high']	a = tief	['deep']
b = breit	['wide']	b = breit	['wide']
c = tief	['deep']	c = ???	

(54) Polish

a = wysoki	['high']	a = głęboki	['deep']
b = długi	['long']	b = długi	['long']
c = szeroki	['wide']	c = szeroki	['wide']

(55) Slovak

a = vysoký	['high']	a = hlboký	['deep']
b = dlhý	['long'] /široký ['wide']	b = dlhý	['long']
c = široký	['wide'] /hlboký ['deep']	c = široký	['wide']



[10] I dare say that English works the same way. But as my informants are still quarreling on whether or not Britannia rules the gaps, I shall not quote English step data before being officially notified of the results.

The gap in (50) is caused by two independent but converging factors. First, as will be recalled, German *tief* covers OBS and OBS-VERT, the latter being the basis for the 'high' - ' deep' alternation we observe in (50), the former causing an interference of (O-1) and (O-3) w.r.t. *tief*. This is the language-particular factor.[10] Second, the Uniqueness Constraint (23) is operative and thus prevents *tief* from occurring twice in (50) despite the fact that it would refer to distinct extensions. This is the universal factor which expectedly prevails.

Finally, (51) shows the observer-based subset of Korean DA terms in two versions. How does this subset score? (O-1) yes (that is the gist of the *selo-kalo* correlation), (O-2) + (O-3) yes (step is in class C), (O-4) no (verticality assignment is separate from maximality assignment). Why are there no ambiguities or gaps? Well, unlike in German, in Korean (O-3) does not interfere with (O-1) as the 'high' - ' deep' alternation [*nophi - kiphi*] is independent of the assignment of *selo-kalo* . So no ambiguity arises nor is there room for a gap.

4.2.2 Finer-grained typology. Another set of staircase data is that in (52) - (55). The point here is the variations we find among languages that are near cognates in i.a. using terms which etymologically share the same common Slavonic origin. Russian, as shown in (48), is high on the scale of proportion-basedness, whereas Polish and Slovak, while using the same adjectival terms, display a tendency towards observer-based assignments — presumably under the influence of German.

(52) Russian

a = vysokij	['high']		
b = dlinnyj	['long']		
c = shirokij	['wide']	/*glubokij	['deep']

(53) German

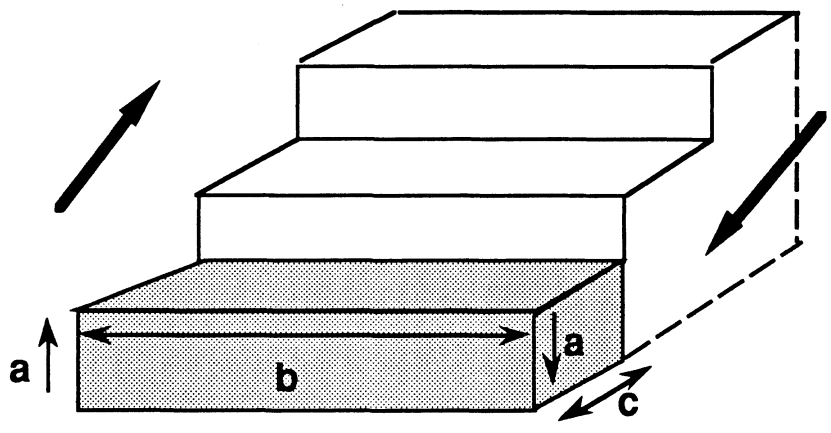
a = hoch	['high']	a = tief	['deep']
b = breit	['wide']	b = breit	['wide']
c = tief	['deep']	c = ???	

(54) Polish

a = wysoki	['high']	a = głęboki	['deep']
b = długi	['long']	b = długi	['long']
c = szeroki	['wide']	c = szeroki	['wide']

(55) Slovak

a = vysoký	['high']	a = hlboký	['deep']
b = dlhý	['long'] /široký ['wide']	b = dlhý	['long']
c = široký	['wide'] /hlboký ['deep']	c = široký	['wide']



[10] I dare say that English works the same way. But as my informants are still quarreling on whether or not Britannia rules the gaps, I shall not quote English step data before being officially notified of the results.

Let's look at how Polish and Slovak stepwise deviate from Russian and approach German. This can best be shown by the following table:

(56)	Mandarin	P-1	P-2	P-3	P-4	
	Russian	P-1	P-2	P-3	P-4	O-1
	Polish	P-1	P-2	O-3	P-4	O-1
	Slovak	P-1	O-2	O-3	P-4	O-1
	German	P-1	O-2	O-3	O-4	O-1

In Mandarin, only Ps apply. In Russian, as we observed in (49), (P-1) - (P-4) apply supplemented by (O-1). Now, (54) shows that Polish has replaced (P-3) with (O-3) - thus allowing for a 'high' - 'deep' alternation (*wysoki* ['high'] – *głęboki* ['deep']), the rest remaining unchanged.[11] Slovak has replaced (P-3) with (O-3) too, but in addition it has replaced (P-2) with (O-2). This is a decisive step towards observer-basedness. Semantically, this is to say that while Russian *glubokij* ['deep'] maintains the selectional restriction to concave or hollow objects, Slovak *hlbokÝ* ['deep'] has loosened it to the effect that the class C of objects to which (O-3) applies resembles that in German.

Note that this is a subtle but interesting case of interference that has gone unnoticed so far. The explanation sketched here seems quite plausible. Despite the fact that in any L I checked so far, DA terms belong to the core lexicon (in this case to the common Slavonic core lexicon), which is normally resistant to lexical or other familiar types of borrowing, they can nevertheless be subject to interference. So, while keeping its place in the common Slavonic core lexicon, the field of Slovak DA terms has undergone subtle changes as to the lexical coverage of DA parameters. The replacement of common Slavonic (P-2) with (O-2) - supposedly under German influence - is a case in point. The subtlety of this change is neatly accounted for by being attributed to the loosening of a selectional restriction of a particular term - as seems to be the case with Slovak *hlbokÝ* ['deep'].

Moreover, also the pragmatic underpinning of this change sounds quite reasonable. Due to geographic, political, and cultural reasons, the West Slavonic languages were and still are in close contact with German. The loosening of the selectional restriction on a term such as *hlbokÝ* ['deep'] is induced by widening the class C of objects to which *hlbokÝ* ['deep'] becomes applicable. This, in turn, is a sort of process that is set going or at least facilitated by the culturally and linguistically intertwined everyday communication that has been charcteristic for this area for centuries.

Unlike sortal constraints, which are a matter to be accounted for on the Conceptual Level, selectional restrictions have to be accounted for on the Semantic Level as they are linked with lexical packaging and hence are subject to language-particular variation (see Lang 1994). So I would like to suggest to enrol selectional restrictions in the list of typologically relevant parameters.

---

[11] As regards the selectional restrictions on *głęboki* ['deep'], Polish seems to be in a transitory state where (P-3) and (O-3) compete with each other. So my informants accepted assigning *głęboki* to e.g. a cupboard "only if it is open".

4.2.3 Inferences. The reconstruction in (56) of the data (52) - (55) is reflected (and hence confirmed) by looking at the sets of inferences in (57) - (59) that draw on the inheritance of dimension assignments from parts to wholes. (57) shows that height if referring to the canonical vertical axis of an object is both part-whole and whole-part inheritable. This is but another aspect of the dominance of the Vertical mentioned in (5) and of the universal principle in (22), so we would not expect any variations across languages as regards the validity of this inference.

(57) (a) The height of the staircase is composed of the heights of the steps
(b) Die Höhe der Treppe ergibt sich aus den Höhen der Stufen [German]
(c) Vysota lestnicy sootvetstvuyet summe vysoty stupenej [Russian]
(d) Wysokość schodów wynika z wysokości schodków [Polish]
(e) Výška schodiště sa skladá zo súčtu výšiek schodov [Slovak]

The inheritability of depth assignments is subject to different conditions. Due to physical unavailability, primary depth assignments - irrespective of covering OBS or OBS-VERT - are hardly inheritable. So if depth enters part-whole inheritance at all, it is confined to cases of OBS-VERT emerging as contextual specification of the vertical. The OBS covered by (53) c = *tief* is not inheritable to a staircase as a whole either. So the staircase talked about in (58) should be one leading downwards into the cellar. With this proviso, the data in (58) seem to prove that depth assignments are part-whole inheritable only in languages where (O-3) applies (hence * for Russian):

(58) (a) The depth of the staircase is composed of the depth of the stairs
(b) Die Tiefe der Treppe ergibt sich aus den Tiefen der Stufen [German]
(c) *Glubina lestnicy sootvetstvuyet summe glubiny stupenej [Russian]
(d) Głębokość schodów wynika z głębokości schodków [Polish]
(e) Hl'bka schodiště sa skladá zo súčtu hl'bok schodov [Slovak]

Things are still different with 'length' (MAX) and 'width' (ACROSS) assignments, which - for the class of objects a staircase belongs to - are not part-whole inheritable. For languages that adhere to proportion-basedness, we will expect an obligatory change of terms, e.g. when stating that the 'length' of the steps equals the 'width' of the staircase. That is what we observe in Russian - cf. (59)(c). In languages like English and German, the observer-based strategy (O-1) - (O-4) prevails due to being applicable to both steps and staircase and thus making a step's width inheritable to the staircase - cf. (59)(a,b). What about the West Slavonic languages in between ?

(59) (a) The width of the steps is equal to the width of the staircase
(b) Die Breite der Stufen ist gleich der Breite der Treppe [German]
(c) Dlina stupenej sootvetstvuyet shirine lestnicy [Russian]
Dlina lestnicy sootvetstvuyet summe shiriny stupenej
*Shirina stupenej sootvetstvuyet shirine lestnicy
(d) Długość schodków odpowiada szerokości schodów [Polish]
Długość schodów odpowiada sumie szerokości schodków
*Szerokość schodków odpowiada szerokości schodów
(e) Dl' žka schodov udáva šírku schodiště [Slovak]
?Šírka schodov udáva šírku schodiště