

Spatial Structuring in Spoken and Signed Language

LEONARD TALMY

University at Buffalo, State University of New York

0. Introduction¹

Linguistic research to date has determined many of the factors that structure the spatial schemas found across spoken languages (e.g. Gruber 1965, Fillmore 1968, Leech 1969, Clark 1973, Bennett 1975, Herskovits 1982, Jackendoff 1983, Zubin and Svorou 1984, as well as myself, Talmy 1983, 2000a, 2000b). It is now feasible to integrate these factors and to determine the comprehensive system they constitute for spatial structuring in spoken language. This system is characterized by several features. With respect to constituency, There is a relatively closed universally available inventory of fundamental spatial elements that in combination form whole schemas. There is a relatively closed set of categories that these elements appear in. And there is a relatively closed small number of particular elements in each category, hence, of spatial distinctions that each category can ever mark. With respect to synthesis, selected elements of the inventory are combined in specific arrangements to make up the whole schemas represented by closed-class spatial forms. Each such whole schema that a closed-class form represents is thus a “pre-packaged” bundling together of certain elements in a particular arrangement. Each language has in its lexicon a relatively closed set of such pre-packaged schemas (larger than that of spatial closed-class forms, due to polysemy) that a speaker must select among in depicting a spatial scene. Finally, with respect to the whole schemas themselves, these schemas can undergo a certain set of processes that extend or deform them. Such processes are perhaps part of the overall system so that a language’s relatively closed set of spatial schemas can fit more spatial scenes.

An examination of signed language² shows that its structural representation of space systematically differs from that in spoken language in the direction of what

¹ An expanded and more updated version of the present paper will appear in Talmy (in press).

² I here approach signed language from the perspective of spoken language because it is not at this point an area of my expertise. For their help with my questions on signed language, my thanks to Paul Dudis, Karen Emmorey, Samuel Hawk, Nini Hoiting, Marlon Kuntze, Scott Liddell, Stephen McCullough, Dan Slobin, Ted Suppala, Alyssa Wolf, and others, - who are not responsible for my errors and oversights.

appear to be the structural characteristics of scene parsing in visual perception. Such differences include the following: Signed language can mark finer spatial distinctions with its inventory of more structural elements, more categories, and more elements per category. It represents many more of these distinctions in any particular expression. It also represents these distinctions independently in the expression, not bundled together into pre-packaged schemas. And its spatial representations are largely iconic with visible spatial characteristics. When formal linguistic investigation of signed language began several decades ago, it was important to establish in the context of that time that signed language was in fact a full genuine language, and the way to do this, it seemed, was to show that it fit the prevailing model of language, the Chomskyan-Fodorian language module. Since then, however, evidence has been steadily accruing that signed language does diverge in various respects from spoken language. The modern response to such observations - far from once again calling into question whether signed language is a genuine language - should be to rethink what the general nature of language is. Our findings suggest that instead of some discrete whole-language module, spoken language and signed language are both based on some more limited core linguistic system that then connects with different further subsystems for the full functioning of the two different language modalities.

1. Fundamental Space-Structuring Elements and Categories in Spoken Language

An initial main finding emerges from analysis of the spatial schemas expressed by closed-class (grammatical) forms across spoken languages. There is a relatively closed and universally available inventory of fundamental conceptual elements that recombine in various patterns to constitute those spatial schemas. These elements fall within a relatively closed set of categories, with a relatively closed small number of elements per category.

1.1. The Target of Analysis

As background to this finding, spoken languages universally exhibit two different subsystems of meaning-bearing forms. One is the “open-class” or “lexical” subsystem, comprised of elements that are great in number and readily augmented - typically, the roots of nouns, verbs, and adjectives. The other is the “closed-class” or “grammatical” subsystem, consisting of forms that are relatively few in number and difficult to augment - including such bound forms as inflections and such free forms as prepositions and conjunctions. As argued in Talmy (2000a, ch. 1), these subsystems basically perform two different functions: open-class forms largely contribute conceptual content, while closed-class forms determine conceptual structure. Accordingly, our discussion focuses on the spatial schemas represented by closed-class forms so as to examine the concepts used by language for structuring purposes.

Across spoken languages, only a portion of the closed-class subsystem regularly represents spatial schemas. We can identify the types of closed-class forms

in this portion and group them according to their kind of schema. The types of closed-class forms with schemas for paths or sites include the following: (1) forms in construction with a nominal, such as prepositions like English *across* (as in *across the field*) or noun affixes like the Finnish illative suffix *-:n* ‘into’, as well as prepositional complexes such as English *in front of* or Japanese constructions with a “locative noun” like *ue* ‘top surface’, (as in *teeburu no ue ni* ‘table GEN top at’ = “on the table”); (2) forms in construction with a verb, such as verb satellites like English *out*, *back* and *apart* (as in *They ran out / back / apart*); (3) deictic determiners and adverbs such as English *this* and *here*; (4) indefinites, interrogatives, relatives, etc., such as English *everywhere* / *whither* / *wherever*; (5) qualifiers such as English *way* and *right* (as in *It’s way / right up there*); and (6) adverbials like English *home* (as in *She isn’t home*). Types of closed-class forms with schemas for the spatial structure of objects include the following: (1) forms modifying nominals such as markers for plexity or state of boundedness, like English *-s* for multiplexing (as in *birds*) or *-ery* for debounding (as in *shrubbery*); (2) numeral classifiers like Korean *chang* ‘planar object’; and (3) forms in construction with the verb, such as some Atsugewi Cause prefixes, like *cu-* ‘as the result of a linear object moving axially into the Figure’. Finally, sets of closed-class forms that represent a particular component of a spatial event of motion/location include the following: (1) the Atsugewi verb-prefix set that represents different Figures; (2) the Atsugewi verb-suffix set that represents different Grounds (together with Paths); (3) the Atsugewi verb-prefix set that represents different Causes; and (4) the Nez Perce verb-prefix set that represents different Manners.

1.2. Determining the Elements and Categories

A particular methodology is used to determine fundamental spatial elements in language. One starts with any closed-class spatial morpheme in any language, considering the full schema that it expresses and a spatial scene that it can apply to. One then determines any factor one can change in the scene so that the morpheme no longer applies to it. Each such factor must therefore correspond to an essential element in the morpheme’s schema. To illustrate, consider the English preposition *across* and the scene it refers to in *The board lay across the road*. Let us here grant the first two elements in the *across* schema (demonstrated elsewhere): (1) a Figure object (here, the board) is spatially related to a Ground object (here, the road); and (2) the Ground is ribbonal - a plane with two roughly parallel line edges that are as long as or longer than the distance between them. The remaining elements can then be readily demonstrated by the methodology. Thus, a third element is that the Figure is linear, generally bounded at both ends. If the board were instead replaced by a planar object, say, some wall siding, one could no longer use the original *across* preposition but would have to switch to the schematic domain of another preposition, that of *over*, as in *The wall siding lay over the road*. A fourth element is that the axes of the Figure and of the Ground are roughly perpendicular. If the board were instead aligned with the road, one

could no longer use the original *across* preposition but would again have to switch to another preposition, *along*, as in *The board lay along the road*. Additionally, a fifth element of the *across* schema is that the Figure is parallel to the plane of the Ground. In the referent scene, if the board were tilted away from parallel, one would have to switch to some other locution such as *The board stuck into / out of the road*. A sixth element is that the Figure is adjacent to the plane of the Ground. If the board were lowered or raised away from adjacency, even while retaining the remaining spatial relations, one would need to switch to locutions like *The board lay (buried) in the road*. / *The board was (suspended) above the road*. A seventh element is that the Figure's length is at least as great as the Ground's width. If the board were replaced by something shorter, for example, a baguette, while leaving the remaining spatial relations intact, one would have to switch from *across* to *on*, as in *The baguette lay on the road*. An eighth element is that the Figure touches both edges of the Ground. If the board in the example retained all its preceding spatial properties but were shifted axially, one would have to switch to some locution like *One end of the board lay over one edge of the road*. Finally, a ninth element is that the axis of the Figure is horizontal (the plane of the Ground is typically, but not necessarily, horizontal). Thus, if one changes the original scene to that of a spear hanging on a wall, one can use *across* if the spear is horizontal, but not if it is vertical, as in *The spear hung across the wall*. / *The spear hung up and down on the wall*. Thus, from this single example, the methodology shows that at least the following elements figure in closed-class spatial schemas: a Figure and a Ground, a point, a line, a plane, a boundary (a point as boundary to a line, a line as boundary to a plane), parallelness, perpendicularity, horizontality, adjacency (contact), and relative magnitude.

In the procedure of systematically testing candidate factors for their relevance, the elements just listed have proved to be essential to the selected schema and hence, to be in the inventory of fundamental spatial elements. But it is equally necessary to note candidates that do not prove out, so as to know which potential spatial elements do not serve a structuring function in language. In the case of *across*, for example, one can probe whether the Figure, like the board in the referent scene, must be planar - rather than simply linear - and coplanar with the plane of the Ground. It can be seen, though, that this is not an essential element to the *across* schema, since this factor can be altered in the scene by standing the board on edge without any need to alter the preposition, as in *The board lay flat / stood on edge across the railway bed*. Thus, coplanarity is not shown by *across* to be a fundamental spatial element. However, it does prove to be so in other schemas, and so in the end must be included in the inventory. This is seen for one of the schemas represented by English *over*, as in *The tapestry hung over the wall*. Here, both the Figure and Ground must be planes and coplanar with each other. If the tapestry here were changed to something linear, say, a string of beads, it is no longer appropriate to use *over* but only something like *against*, as in *The string of beads hung *over / against the wall*. Now, another candidate element - that the Figure must be rigid, like the board in the scene - can be tested and again found to

be inessential to the *across* schema, since a flexible linear object can be substituted for the board without any need to change the preposition, as seen in *The board / The cable lay across the railway bed*. Here, however, checking this candidate factor across numerous spatial schemas in many languages might well never yield a case in which it does figure as an essential element and so would be kept off the inventory.

This methodology affords a kind of existence proof: it can demonstrate that some element does occur in the universally available inventory of structural spatial elements since it can be seen to occur in at least one closed-class spatial schema in at least one language. The procedure is repeated numerous times across many languages to build up a sizable inventory of elements essential to spatial schemas.

The next step is to discern whether the uncovered elements comprise particular structural categories and, if so, to determine what these categories are. It can be observed that for certain sets of elements, the elements in a set are mutually incompatible - only one of them can apply at a time at some point in a schema. Such sets are here taken to be basic spatial categories. Along with their members, such categories are also part of language's fundamental conceptual structuring system for space. A representative sample of these categories is presented next.

It will be seen that these categories generally have a relatively small membership. This finding depends in part on the following methodological principles. An element proposed for the inventory should be as coarse-grained as possible - that is, no more specific than is warranted by cross-schema analysis. Correlatively, in establishing a category, care must be taken that it include only the most generic elements that have actually been determined - that is, that its membership have no finer granularity than is warranted by the element-abstraction procedure. For example, the principle of mutual incompatibility yields a spatial category of "relative orientation" between two lines or planes, a category with perhaps only two member elements (both already seen in the *across* schema): approximately parallel and approximately perpendicular. Some evidence additionally suggests an intermediary "oblique" element as a third member of the category. Thus, some English speakers may distinguish a more perpendicular sense from a more oblique sense, respectively, for the two verb satellites *out* and *off*, as in *A secondary pipe branches out / off from the main sewer line*. In any case, though, the category would have no more than these two or three members. Although finer degrees of relative orientation can be distinguished by other cognitive systems, say, in visual perception and in motor control, the conceptual structuring subsystem of language does not include anything finer than the two- or three-way distinction. The procedures of schema analysis and cross-schema comparison, together with the methodological principles of maximum granularity for elements and for category membership, can lead to a determination of the number of structurally distinguished elements ever used in language for a spatial category.

1.3. Sample Categories and their Member Elements

The fundamental categories of spatial structure in the closed-class subsystem of spoken language fall into three classes according to the aspect of a spatial scene they pertain to: the segmentation of the scene into individual components, the properties of an individual component, and the relations of one such component to another. In a fourth class are categories of nongeometric elements frequently found in association with spatial schemas. A sampling of categories and their member elements from each of these four classes is presented next. The examples provided here are primarily drawn from English but can be readily multiplied across a diverse range of languages (see Talmy 2000a, ch. 3).

1.3.1. Categories Pertaining to Scene Segmentation

The class designated as scene segmentation may include only one category, that of “major components of a scene”, and this category may contain only three member elements: the Figure, the Ground, and a secondary Reference Object. Figure and Ground were already seen for the *across* schema. Schema comparison shows the need to recognize a third scene component, the Secondary Reference Object - in fact, two forms of it: encompassive of or external to the Figure and Ground. The English preposition *near*, as in *The lamp is near the TV* specifies the location of the Figure (the lamp) only with respect to the Ground (the TV). But localizing the Figure with the preposition *above*, as in *The lamp is above the TV*, requires knowledge not only of where the Ground object is, but also of the encompassive earth-based spatial grid, in particular, of its vertical orientation. Thus, *above* requires recognizing three components within a spatial scene, a Figure, a Ground, and a Secondary Reference Object of the encompassive type. Comparably, the schema of the *past* in *John is past the border* only relates John as Figure to the border as Ground. One could say this sentence on viewing the event through binoculars from either side of the border. But *John is beyond the border* can be said only by someone on the side of the border opposite John, hence the *beyond* schema establishes a perspective point at that location as a secondary Reference Object - in this case, of the external type.

1.3.2. Categories Pertaining to an Individual Scene Component

A number of categories pertain to the characteristics of an individual spatial scene component. This is usually one of the three major components resulting from scene segmentation - the Figure, Ground, or Secondary Reference Object - but it could be others, such as the path line formed by a moving Figure. One such category is that of “dimension” with four member elements: zero dimensions for a point, one for a line, two for a plane, and three for a volume. Some English prepositions require a Ground object schematizable for only one of the four dimensional possibilities. Thus, the schema of the preposition *near* as in *near the dot* requires only that the Ground object be schematizable as a point. *Along*, as in *along the trail*, requires that the Ground object be linear. *Over* as in *a tapestry over a wall* requires a planar Ground. And *throughout*, as in *cherries throughout*

the jello, requires a volumetric Ground. A second category is that of “number” with perhaps four members: one, two, several, and many. Some English prepositions require a Ground comprising objects in one or another of these numbers. Thus, *near* requires a Ground consisting of just one object, *between* of two objects, *among* of several objects, and *amidst* of numerous objects, as in *The basketball lay near the boulder / between the boulders / among the boulders / amidst the cornstalks*. The category of number appears to lack any further members - that is, closed-class spatial schemas in languages around the world seem never to incorporate any other number specifications - such as ‘three’ or ‘even-numbered’ or ‘too many’. A third category is that of “motive state”, with two members: motion and stationariness. Several English prepositions mark this distinction for the Figure. Thus, in one of its senses, *at* requires a stationary Figure, as in *I stayed / *went at the library*, while *into* requires a moving Figure, as in *I went / *stayed into the library*. Other prepositions mark this same distinction for the Ground object (in conjunction with a moving Figure). Thus, *up to* requires a stationary Ground (here, the deer), as in *The lion ran up to the deer*, while *after* requires a moving Ground as in *The lion ran after the deer*. Apparently no spatial schemas mark such additional distinctions as motion at a fast vs. slow rate, or being located at rest vs. remaining located fixedly. A fourth category is that of “state of boundedness” with two members: bounded and unbounded. The English preposition *along* requires that the path of a moving Figure be unbounded, as shown by its compatibility with a temporal phrase in *for* but not *in*, as in *I walked along the pier for 10 minutes / *in 20 minutes*. But the spatial locution *the length of* requires a bounded path, as in *I walked the length of the pier in 20 Minutes / *for 10 minutes*. While some spatial schemas have the bounded element at one end of a line and the unbounded element at the other end, apparently no spatial schema marks any distinctions other than the two cited states of boundedness, such as a cline of gradually increasing boundedness along a line.

Continuing the sampling of this class, a fifth category is that of “directedness” with two members: basic and reversed. A schema can require one or the other of these elements for an encompassive Ground object, as seen for the English prepositions in *The axon grew along / against the chemical gradient*, or for the Atsugewi verb satellites for (moving) ‘downstream’ and ‘upstream’. Or it can require one of the member elements for an encompassive Secondary Reference Object (here, the line), as in *Mary is ahead of / behind John in line*. A sixth category is “type of geometry” with two members: rectilinear and radial. This category can apply to an encompassive Secondary Reference Object to yield reference frames of the two geometric types. Thus, in a subtle effect, the English verb satellite *away*, as in *The boat drifted further and further away / out from the island*, tends to suggest a rectilinear reference frame in which one might picture the boat moving rightward along a corridor or sea lane with the island on the left (as if along the x-axis of a Cartesian grid). But *out* tends to suggest a radial reference frame in which the boat is seen moving from a center point along a radius through a continuum of concentric circles. The radial member of the

geometry category can involve motion about a center, along a radius, or along a periphery. The first of these is the basis for a further category, that of “orientation of spin axis”, with two members: vertical and horizontal. The English verb satellites *around* and *over* specify motion of the Figure about a vertical or horizontal spin axis, respectively, as in *The pole spun around / toppled over* and in *I turned the pail around / over*. An eighth category is “phase of matter”, with three main members, solid, liquid, and empty space, and perhaps a fourth member, fire. Thus, among the dozen or so Atsugewi verb satellites that subdivide the semantic range of English *into* plus a Ground object, the suffix *-ik’s* specifies motion horizontally into solid matter (as chopping an ax into a tree trunk), *-ic’t* specifies motion into liquid, *-ipsnu* specifies motion into the empty space of a volumetric enclosure, and *-caw* specifies motion into a fire. The phase of matter category even figures in some English prepositions, albeit covertly. Thus, *in* can apply to a Ground object of any phase of matter, whereas *inside* can apply only to one with empty space, as seen in *The rock is in / inside the box; in / *inside the ground; in / *inside the puddle of water; in / *inside the fire*. A final category in this sampled series is that of “state of consolidation” with apparently two members: compact (precisional) and diffuse (approximative). The English locative prepositions *at* and *around* distinguish these two concepts, respectively, for the area surrounding a Ground object, as in *The other hiker will be waiting for you at / around the landmark*. The same distinction is marked by the two deictic adverbs in *The hiker will be waiting for you there / thereabouts*. In addition to this sampling, some ten or so further categories pertaining to properties of an individual schema component, each category with a small number of fixed contrasts, can be readily identified.

1.3.3. Categories Pertaining to the Relation of One Scene Component to Another

Another class of categories pertains to the relations that one scene component can bear to another. One such category was described earlier, that of “relative orientation”, with two or three members: parallel, perpendicular, and perhaps oblique. A second such category is that of “degree of remove”, of one scene component from another. This category appears to have four or five members, two with contact between the components - coincidence and adjacency – and two or three without contact - proximal, perhaps medial, and distal remove. Some pairwise contrasts in English reveal one or another of these member elements for a Figure relating to a Ground. Thus, the locution *in the front of*, as in *The carousel is in the front of the fairground*, expresses coincidence, since the carousel as Figure is represented as being located in a *part* of the fairground as Ground. But *in front of* (without a *the*) as in *The carousel is in front of the fairground*, indicates proximality, since the carousel is now located outside the fairground and near it but not touching it. The distinction between proximal and distal can be teased out by noting that *in front of* can only represent a proximal but not a distal degree of remove, as seen in that one can say *The carousel is 20 feet in front of the fairground*, but not, **The*

carousel is 20 miles in front of the fairground, whereas *above* allows both proximal and distal degrees of remove, as seen in *The hawk is 1 foot / 1 mile above the table*. The distinction between adjacency and proximality is shown by the prepositions *on* and *over*, as in *The fly is on / over the table*. Need for a fifth category member of ‘medial degree of remove’ might come from languages with a ‘here / there / yonder’ kind of distinction in their deictic adverbs or demonstratives. A third category in this series is that of “degree of dispersion” with two members: sparse and dense. To begin with, English can represent a set of multiple Figures, say, 0-dimensional peas, as adjacent to or coincident with a 1-, 2-, or 3-dimensional Ground, say, with a knife, a tabletop, or aspic, in a way neutral to the presence or absence of dispersion, as in *There are peas on the knife; on the table; in the aspic*. But in representing dispersion as present, English can (or must) indicate its degree. Thus, a sparse degree of dispersion is indicated by the addition of the locution *here and there*, optionally together with certain preposition shifts, as in *There are peas here and there on / along the knife; on / over the table; in the aspic*. And for a dense degree of dispersion, English has the three specialized forms *all along*, *all over* and *throughout*, as seen in *There are peas all along the knife; all over the table; throughout the aspic*. A fourth category is that of “path contour” with perhaps some four members: straight, arced, circular, and meandering. Some English prepositions require one or another of these contour elements for the path of a Figure moving relative to a Ground. Thus, *across* indicates a straight path, as seen in *I drove across the plateau / *hill*, while *over* - in its usage referring to a single path line - indicates an arced contour, as in *I drove over the hill / *plateau*. In one of its senses, *around* indicates a roughly circular path, as in *I walked around the maypole*, and *about* indicates a meandering contour, as in *I walked about the town*. Some ten or so additional categories for relating one scene component to another, again each with its own small number of member contrasts, can be readily identified.

1.3.4. Nongeometric Categories

All the preceding elements and their categories have broadly involved geometric characteristics of spatial scenes or the objects within them - that is, they have been genuinely spatial. But a number of nongeometric elements are recurrently found in association with otherwise geometric schemas. One category of such elements is that of “force dynamics” (see Talmy 2000a, ch. 7) with two members: present and absent. Thus, geometrically, the English prepositions *on* and *against* both represent a Figure in adjacent contact with a Ground, but in addition, *on* indicates that the Figure is supported against the pull of gravity through that contact while *against* indicates that it is not, as seen in *The poster is on / *against the wall* and *The floating helium balloon is against / *on the wall*. A second nongeometric category is that of “accompanying cognitive/affective state”, though its extent of membership is not clear. One recurrent member, however, is the attitude toward something of its being unknown, mysterious, or risky. Perhaps in combination with elements of inaccessibility or nonvisibility, this category member is associ-

ated with the Figure's location in the otherwise spatial indications of the English preposition *beyond*, whereas it is absent from the parallel locution *on the other side of*, as in *He is beyond / on the other side of the border* (both these locutions - unlike *past* seen above - are otherwise equivalent in establishing a viewpoint location as an external Secondary Reference Object). A third nongeometric category, - in the class that relates one scene component to another - is that of "relative priority", with two members: coequal and main/ancillary. The English verb satellites *together* and *along* both indicate joint participation, as seen in *I jog together / along with him*. But *together* indicates that the Figure and the Ground are coequal partners in the activity, whereas *along* indicates that the Figure entity is ancillary to the Ground entity, who would be assumed to engage in the activity even if alone (see Talmy 2000b, ch. 3).

1.4. Properties of the Inventory

By our methodology, the universally available inventory of structural spatial elements includes all elements that appear in at least one closed-class spatial schema in at least one language. These elements may indeed be equivalent in their sheer availability for use in schemas. But beyond that, they appear to differ in their frequency of occurrence across schemas and languages, ranging from very common to very rare. Accordingly, the inventory of elements - and perhaps also that of categories - may have the property of being hierarchical, with entries running from the most to the least frequent. Such a hierarchy suggests asking whether the elements in the inventory, the categories in the inventory, and the elements in each category form fully closed memberships. That is, does the hierarchy end at a sharp lower boundary or trail off indefinitely? With many schemas and languages already examined, our sampling method may have yielded all the commoner elements and categories, but as the process slows down in the discovery of the rarer forms, will it asymptotically approach some complete constituency and distinctional limit in the inventory, or will it be able to go on uncovering sporadic novel forms as they develop in the course of language change?

The latter seems likelier. Exotic elements with perhaps unique occurrence in one or a few schemas in just one language can be noted, including in English. Thus, in referring to location at the interior of a wholly or partly enclosed vehicle, the prepositions *in* and *on* distinguish whether the vehicle lacks or possesses a walkway. Thus, one is *in* a car but *on* a bus, *in* a helicopter but *on* a plane, *in* a grain car but *on* a train, and *in* a rowboat but *on* a ship. Further, Fillmore has observed that this *on* also requires that the vehicle be currently in use as transport: *The children were playing in / *on the abandoned bus in the junkyard*. Thus, schema analysis in English reveals the element '(partly) enclosed vehicle with a walkway currently in use as transport'. This is surely one of the rarer elements in schemas around the world, perhaps unique, and its existence, along with that of various others that can be found, suggests that indefinitely many more of them can sporadically arise.

In addition to being only relatively closed at its hierarchically lower end, the inventory may include some categories whose membership seems not to settle down to a small fixed set. One such category may be that of “intrinsic parts”. Frequently encountered are the five member elements ‘front’, ‘side’, ‘back’, ‘top’, and ‘bottom’, as found in the English prepositions in *The cat lay before / beside / behind / atop / beneath the TV*. But languages like Mixtec seem to distinguish a rather different set of intrinsic parts in its spatial schemas, while Makah (Matthew Davidson, personal communication) distinguishes many more and finer parts, such as with its verb suffixes for ‘at the ankle’ and ‘at the groin’.

Apart from any fuzzy lower boundary and noncoalescing categories, there does appear to exist a graduated inventory of basic spatial elements and categories that is universally available and, in particular, is relatively closed. Bowerman (e.g. 1989) has raised the main challenge to this notion. She notes, for example, that at the same time that children acquiring English learn its *in/on* distinction, children acquiring Korean learn its distinction between *kkita* ‘put [Figure] in a snug fit with [Ground]’ and *nehta* ‘put [Figure] in a loose fit with [Ground]’ she argues that since the elements ‘snug fit’ and ‘loose fit’ are presumably rare among spatial schemas across languages, they do not come from any preset inventory, one that might plausibly be innate, but rather are learned from the open-ended semantics of the adult language. My reply is that the spatial schemas of genuinely closed-class forms in Korean may well still be built from the proposed inventory elements, and that the forms she cites are actually open verbs. Open-class semantics - whether for space or other domains - seems to involve a different cognitive subsystem, drawing from finer discriminations within a broader perceptual / conceptual sphere. The Korean verbs are perhaps learned at the same age as English space-related open-class verbs like *squeeze*. Thus, English-acquiring children probably understand that *squeeze* involves centripetal pressure from encircling or bi-/multi-laterally placed Antagonists (typically the arm(s) or hand(s)) against an Agonist that resists the pressure but yields down to some smaller compass where it blocks further pressure, and hence that one can squeeze a teddy bear, a tube of toothpaste, or a rubber ball, but not a piece of string or sheet of paper, juice or sugar or the air, a tabletop or the corner of a building. Thus, Bowerman’s challenge may be directed at the wrong target, leaving intact the proposed roughly preset inventory of basic spatial building blocks.

1.5. Basic Elements Assembled into Whole Schemas

The procedure so far has been analytic, starting with the whole spatial schemas expressed by closed-class forms and abstracting from them an inventory of fundamental spatial elements. But the investigation must also include a synthetic procedure: examining the ways in which individual spatial elements are assembled to constitute whole schemas. Something of such an assembly was implicit in the initial discussion of the *across* schema. But an explicit example here can better illustrate this part of the investigation.

Consider the schema represented by the English preposition *past* as in *The ball sailed past my head at exactly 3 PM*. This schema is built out of the following fundamental spatial elements (from the indicated categories) in the indicated arrangements and relationships. There are two main scene components (members of the “major scene components” category), a Figure and a Ground (here, the ball and my head, respectively). The Figure is schematizable as a 0-dimensional point (a member element of the “dimension” category). This Figure point is moving (a member element of the “motive state” category). Hence it forms a one-dimensional line (a member of the “dimension” category”). This line constitutes the Figure’s “path”. The Ground is also schematizable as a 0-dimensional point (a member of the “dimension” category). There is a point P at a proximal remove (a member of the “degree of remove” category) from the Ground point, forming a 1-dimensional line with it (a member of the “dimension” category). This line is parallel (a member of the “relative orientation” category) to the horizontal plane (a member of the “intrinsic parts” category) of the earth-based grid (a member of the major scene components” category). The Figure’s path is perpendicular (a member of the “relative orientation” category) to this line. The Figure’s path is also parallel to the horizontal plane of the earth-based grid. If the Ground object has a front, side, and back (members of the “intrinsic parts” category), then point P is proximal to the side part. A nonboundary point (a member of the “state of boundedness” category) of the Figure’s path becomes coincident (a member of the “degree of remove” category) with point P at a certain point of time.

Note that here the Figure’s path must be specified as passing through a point proximal to the Ground because if it instead passed through the Ground point, one would switch from the preposition *past* to *into*, as in *The ball sailed into my head*, and if it instead past through some distal point, one might rather say something like *The ball sailed along some ways away from my head*. And the Figure’s path must be specified both as horizontal and as located at the side portion of the Ground because, for example here, if the ball were either falling vertically or traveling horizontally at my front, one could no longer say that it sailed past my head.

The least understood aspect of the present investigation is what well-formedness conditions, if any, may govern the legality of such combinations. As yet, no obvious principles based, say, on geometric simplicity, symmetry, consistency, or the like are seen to control the patterns in which basic elements assemble into whole schemas. On the one hand, some seemingly byzantine combinations - like the schemas seen above for *across* and *past* - occur with some regularity across languages. On the other hand, much simpler combinations seem never to occur as closed-class schemas. For example, one could imagine assembling elements into the following schema: down into a surround that is radially proximal to a center point. One could invent a preposition *apit* to represent this schema, as used in *I poured water apit my house*” to refer to my pouring water down into a nearby hole dug in the field around my house. But such schemas are not found. Similarly, a number of schematic distinctions in, for example, the domain of

rotation are regularly marked by signed languages, as seen below, and could readily be represented with the inventory elements available to spoken languages, yet they largely do not occur. It could be argued that the spoken language schemas are simply the spatial structures most often encountered in everyday activity. But that would not explain why the additional sign-language schemas - presumably also reflective of everyday experience - do not show up in spoken languages. Besides, the different sets of spatial schemas found in different spoken languages are diverse enough from each other that arguing on the basis of the determinative force of everyday experience is problematic. Something else is at work but it is not yet clear what that is.

1.6. Properties and Processes Applying to Whole Spatial Schemas

It was just seen that selected elements of the inventory are combined in specific arrangements to make up the whole schemas represented by closed-class spatial forms. Each such whole schema is thus a “pre-packaged” bundling together of certain elements in a particular arrangement. Each language has in its lexicon a relatively closed set of such pre-packaged schemas - one larger than that of its spatial closed-class forms, because of polysemy. A speaker of the language must select among these schemas in depicting a spatial scene. We now observe that such schemas, though composite, have a certain unitary status in their own right, and that certain quite general properties and processes can apply to them. In particular, certain properties and processes allow a schema represented by a closed-class form to generalize to a whole family of schemas. In the case of a generalizing *property*, all the schemas of a family are of equal priority. On the other hand, a generalizing *process* acts on a schema that is somehow basic, and either extends or deforms it to yield nonbasic schemas. (see Talmy 2000a ch. 1 and 3, 2000b ch. 5). Such properties and processes are perhaps part of the overall spoken-language system so that any language’s relatively closed set of spatial closed-class forms and the schemas that they basically represent can be used to match more spatial structures in a wider range of scenes.

Looking first at generalizing properties of spatial schemas, one such property is that they exhibit a topological or topology-like neutrality to certain factors of Euclidean geometry. Thus, they are magnitude neutral, as seen in such facts as that the *across* schema can apply to a situation of any size, as in *The ant crawled across my palm / The bus drove across the country*. Further, they are largely shape-neutral, as seen by such facts as that, while the *through* schema requires that the Figure form a path with linear extent, it lets that line take any contour, as in *I zig-zagged / circled through the woods*. And they are bulk-neutral, as seen by such facts as that the *along* schema requires a linear Ground without constraint on the Ground’s radial extension, as in *The caterpillar crawled up along the filament / tree trunk*. Thus, while holding to their specific constraints, schemas can vary freely in other respects and so cover a range of spatial configurations.

Among the processes that extend schemas, one is that of “extendability from the prototype”, which can serve as an alternative interpretation for some forms of

neutrality. Thus, in the case of shape, as for the *through* schema above, this schema could alternatively be conceived as prototypically involving a strait path line for the Figure, one that can then be bent to any contour. And, in the case of bulk, as for the *along* schema above, this schema could be thought prototypically to involve a purely 1-dimensional line that then can be radially inflated. Another such process is “extendability in ungoverned dimensions”. By this process, a scene component of dimensionality N in the basic form of a schema can generally be raised in dimensionality to form a line, plane, or volume aligned in a way not conflicting with the schema’s other requirements. To illustrate, it was seen earlier under the “geometric type” category that the English verb satellite *out* has a schema involving a point Figure moving along a radius away from a center point through a continuum of concentric circles, as in *The boat sailed further and further out from the island*. This schema with the Figure idealizable as a point is the basic form. But the same satellite can be used when this Figure point is extended to form a 1-dimensional line along a radius, as in *The caravan of boats sailed further and further out from the island*. And the *out* can again be used if the Figure point were instead extended as a 1-dimensional line forming a concentric circle, as in *A circular ripple spread out from where the pebble fell into the water*. In turn, such a concentric circle could be extended to fill in the interior plane, as in *The oil spread out over the water from where it spilled*. Alternatively, the concentric circle could have been extended in the vertical dimension to form a cylinder, as in *A ring of fire spread out as an advancing wall of flames*. Or again, the circle could have been extended to form a spherical shell, as in *The balloon I blew into slowly puffed out*. And such a shell can be extended to fill in the interior volume, as in *The leavened dough slowly puffed out*. One more process in this set is “extendability across motive states”. A schema basic for one motive state and Figure geometry can in general be systematically extended to another motive state and Figure geometry. For example, a closed-class form whose most basic schema pertains to a point Figure moving to form a path can generally serve as well to represent the related schema with a stationary linear Figure in the same location as the path. Thus, probably the most basic *across* schema is actually for a moving point Figure, as in *The gopher ran across the road*. By the present process, this schema can extend to the static linear Figure schema first seen in *The board lay across the road*. All the spatial properties uncovered for that static schema hold as well for the present basic dynamic schema, which in fact is the schema in which these properties originally arise.

Among the processes that deform a schema, one is that of “stretching”, which allows a slight relaxing of one of the normal constraints. Thus, in the *across* schema, where the Ground plane is either a ribbon with a long and short axis or a square with equal axes, a static linear Figure or the path of a moving point Figure must be aligned with the short Ground axis or with one of its equal axes. Accordingly, one can say *I swam across the canal* and *I swam across the square pool* when moving from one side to the other, but one cannot say **I swam across the canal* when moving from one end to the other. But, by moderately stretching one

axis length relative to the other, one might just about be able to say *I swam across the pool* when moving from one end to the other of an oblong pool. Another schema deforming process is that of “feature cancellation”, in which a particular complex of elements in the basic schema is omitted. Thus, the preposition *across* can be used in *The shopping cart rolled across the boulevard and was hit by an oncoming car*, even though one feature of the schema - ‘terminal point coincides with the distal edge of the Ground ribbon’ - is canceled from the Figure’s path. Further, both this feature and the feature ‘beginning point coincides with the proximal edge of the Ground ribbon’ are canceled in *The tumbleweed rolled across the prairie for an hour*. Thus, the spoken language system includes a number of generalizing properties and processes that allow the otherwise relatively closed set of abstracted or basic schemas represented in the lexicon of any single language to be applicable to a much wider range of spatial configurations.

2. Spatial Structuring in Signed Language

All the preceding findings on the linguistic structuring of space have been based on the patterns found in spoken languages. The inquiry into the fundamental concept structuring system of language leads naturally to investigating its character in another major body of linguistic realization, signed language. The value in extending the inquiry in this way would be to discover whether the spatial structuring system is the same or is different in certain respects across the two language modalities, with either discovery having major consequences for cognitive theory.

In this research extension, a problematic issue is exactly what to compare between spoken and signed language. The two language systems appear to subdivide into somewhat different sets of subsystems. Thus, heuristically, the generalized spoken language system can be thought to consist of an open-class or lexical subsystem (generally representing conceptual content); a closed-class or grammatical subsystem (generally representing conceptual structure); a gradient subsystem of “vocal dynamics” (including loudness, pitch, timbre, rate, distinctness, unit separation); and an accompanying somatic subsystem (including facial expression, gesture, and “body language”). On the other hand, by one provisional proposal, the generalized sign language system might instead divide up into the following: a subsystem of lexical forms (including noun, verb, and adjective signs); an “inflectional” subsystem (including modulations of lexical signs for person, aspect); a subsystem of size-and-shape specifiers (or SASS’s); a subsystem of so-called “classifier constructions”; a gestural subsystem (along a gradient of incorporation into the preceding subsystems); a subsystem of face, head, and torso representations; a gradient subsystem of “bodily dynamics” (including amplitude, rate, distinctness, unit separation); and an associated or overlaid somatic subsystem (including further facial expression and “body language”). In particular here, the subsystem of classifier constructions – which is apparently present in all signed languages - is a formally distinct subsystem dedicated solely to the sche-

matic structural representation of objects moving or located with respect to each other in space (see Liddell forthcoming, Emmorey in press).³

The research program of comparing the representation of spatial structure across the two language modalities ultimately requires considering the two whole systems and all their subsystems. But the initial comparison - the one adopted here - should be between those portions of each system most directly involved with the representation of spatial structure. In spoken language, this is that part of the closed-class subsystem that represents spatial structure and, in signed language, it is the subsystem of classifier constructions. Spelled out, the shared properties that make this initial comparison apt include the following. First, of course, both subsystems represent objects relating to each other in space. Second, in terms of the functional distinction between “structure” and “content” described earlier, each of the subsystems is squarely on the structural side. In fact, analogous structure-content contrasts occur. Thus, the English closed-class form *into* represents the concept of a path that begins outside and ends inside an enclosure in terms of schematic structure, in contrast with the open-class verb *enter* that represents the same concept in terms of substantive content (see Talmy 2000a, ch. 1 for this structure-content distinction). Comparably, any of the formations within a classifier expression for such an outside-to-inside path represents it in terms of its schematic structure, in contrast with the unrelated lexical verb *sign* that can be glossed as ‘enter’. Third, in each subsystem, a schematic structural form within an expression in general can be semantically elaborated by a content form that joins or replaces it within the same expression. Thus, in the English sentence *I drove it (- the motorcycle-) in (to the shed)* the parenthesized forms optionally elaborate on the otherwise schematically represented Figure and Ground. Comparably, in the ASL sentence “(SHED) (MOTORCYCLE) vehicle-move-into-enclosure”, the optionally signed forms within parentheses elaborate on the otherwise schematic Figure and Ground representations within the hyphenated classifier expression.

To illustrate the classifier system, a spatial event that English could express as *The car drove past the tree* could be expressed in ASL as follows: The signer’s dominant hand, used to represent the Figure object, here has a “3 handshape” (index and middle fingers extended forward, thumb up) to represent a land vehicle. The nondominant hand, used to represent the Ground object, here involves an upright “5 handshape” (forearm held upright with the five fingers extended upward and spread apart) to represent a tree. The dominant hand is moved horizontally across the signer’s torso and past the nondominant forearm. Further though, this basic form could be modified or augmented to represent additional particulars of the referent spatial event. Thus, the dominant hand can show additional characteristics of the path. For example, the hand could move along a curved path to indicate that the road being followed was curved, it could

³ The “classifier” label for this subsystem - originally chosen because its constructions usually include a classifier-like handshape - can be misleading. An apter term might be the “Motion-event subsystem”.

slant upward to represent an uphill course, or both could be shown together. The dominant hand can additionally show the manner of the motion. For example, as it moves along, it could oscillate up and down to indicate a bumpy ride, or move quickly to indicate a swift pace, or both could be shown together, as well as with the preceding two path properties. And the dominant hand can show additional relationships of the Figure to the Ground. For example, it could pass nearer or farther from the nondominant hand to indicate the car's distance from the tree when passing it, it could make the approach toward the nondominant hand longer (or shorter) than the trailing portion of the path to represent the comparable relationship between the car's path and the tree, or it could show both of these together or, indeed, with all the preceding additional characteristics.

The essential finding of how signed language differs from spoken language is that it more closely parallels what appear to be the structural characteristics of scene parsing in visual perception. This difference can be observed in two venues, the universally available spatial inventory and the spatial expression. These two venues are discussed next in turn.

2.1. In the Inventory

The inventory of forms for representing spatial structure available to signed language has a greater total number of fundamental elements, a greater number of categories, and generally a greater number of elements per category than the spoken language inventory. More specifically, the classifier subsystem of signed language has many of the same space-structuring categories as in the closed-class subsystem of spoken language, but it also has many categories not present there, whereas spoken language may have no categories that are absent from signed language. Comparing the membership of the corresponding categories in terms of discrete elements, the number of basic elements per category in signed language ranges from being the same as that for spoken language to being very much greater. Further, though, while the membership of some categories in signed language may well consist of discrete elements, that of others appears to be gradient. Here, any procedure of tallying some fixed number of discrete elements in a category must give way to determining the approximate fineness of distinctions that can be practicably made for that category. So while some corresponding categories across the two language modalities may otherwise be quite comparable, their memberships can be of two different types, discrete vs. analog. Altogether, then, given its greater number of categories, generally larger membership per category, and a frequently gradient type of membership, the inventory of forms for building a schematic spatial representation available to the classifier subsystem of signed language is more extensive and finer than for the closed-class subsystem of spoken language. This greater extensiveness and finer granularity of spatial distinctions seems more comparable to that of spatial parsing in visual perception.

The following are some spatial categories in common across the two language modalities, but with increasing disparity in size of membership. First, some

categories appear to be quite comparable across the two modalities. Thus, both the closed-class subsystem of spoken language and the classifier subsystem of signed language structurally segment a scene into the same three components, a Figure, a Ground, and a secondary Reference Object. Both subsystems represent the category of dimensionality with the same four members – a point, a line, a plane, and a volume. And both mark the same two degrees of boundedness: bounded and unbounded.

For certain categories, signed language has just a slightly greater membership than does spoken language. Thus, for motive state, signed language structurally represents not only moving and being located, but also remaining fixedly located - a concept that spoken languages typically represent in verbs but not in their spatial preposition-like forms.

For other spatial categories, signed language has a moderately greater membership than spoken language. In some of these categories, the membership is probably gradient, but without the capacity to represent many fine distinctions clearly. Thus, signed language can apparently mark moderately more degrees of remove than spoken language's four or five members in this category. It can also apparently distinguish moderately more path lengths than the two - short and long - that spoken language marks structurally (as in English *The bug flew right / way up there*). And while spoken language can mark at most three distinctions of relative orientation - parallel, perpendicular, and oblique - signed language can distinguish a moderately greater number, for example, in the elevation of a path's angle above the horizontal, or in the angle of the Figure's axes to that of the Ground (e.g. in the placement of a pole against a wall).

Finally, there are some categories for which signed language has an indefinitely greater membership than spoken language. Thus, while spoken language structurally distinguishes some four path contours as seen in section 2.3.3, signed language can represent perhaps indefinitely many more, including zigzags, spirals, and ricochets. And for the category "locus within referent space", spoken language can structurally distinguish perhaps at most three loci relative to the speaker's location - 'here', 'there', and 'yonder' - whereas sign language can distinguish indefinitely many more within sign space.

Apart from membership differences across common categories, signed language represents some categories not found in spoken language. One such category is the relative lengths of a Figure's path before and after encounter with the Ground. Or again, signed language can represent not only the category of "degree of dispersion" (which spoken language was seen to represent in section 2.3.3), but also the category "pattern of distribution". Thus, in representing multiple Figure objects dispersed over a planar surface, it could in addition structurally indicate that these Figure objects are linear (as with dry spaghetti over a table) and are arrayed in parallel alignment, crisscrossing, or in a jumble.

This difference in the number of structurally marked spatial category and element distinctions between spoken and signed language can be highlighted with a closer analysis of a single spatial domain, that of rotational motion. As seen

earlier, the closed-class subsystem in spoken language basically represents only one category within this domain, that of “orientation of spin axis”, and within this category distinguishes only two member elements, vertical and horizontal. These two member elements are expressed, for example, by the English verb satellites *around* and *over* as in *The pole spun around / toppled over*. ASL, by contrast, distinguishes more degrees of spin axis orientation and, in addition, marks several further categories within the domain of rotation. Thus, it represents the category of “amount of rotation” and within this category can readily distinguish, say, whether the arc of a Figure’s path is less than, exactly, more than, or many times one full circuit. These are differences that English might offer for inference only from the time signature, as in *I ran around the house for 20 seconds / in 1 minute / for 2 minutes / for hours*, while using the same single spatial form *around* for all these cases. Further, while English would continue using just *around* and *over*, ASL further represents the category of “relation of the spin axis to an object’s geometry” and marks many distinctions within this category. Thus, it can structurally mark the spin axis as being located at the center of the turning object - as well as whether this object is planar like a CD disk, linear like a propeller, or an aligned cylinder like a pencil spinning on its point. It distinguishes this from the spin axis located at the boundary of the object - as well as whether the object is linear like the “hammer” swung around in a hammer toss, a transverse plane like a swinging gate, or a parallel plane like a swung cape. And it further distinguishes these from the spin axis located at a point external to the object - as well as whether the object is point-like like the earth around the sun, or linear like a spinning hoop. Finally, ASL can structurally represent the category of “uniformity of rotation” with its two member elements, uniform and nonuniform, where English could mark this distinction only with an open-class form, like the verbs in *The hanging rope spun / twisted around*, while once again continuing with the same single structural closed-class form *around*. Thus, while spoken language structurally marks only a minimal distinction of spin axis orientation throughout all these geometrically distinct forms of rotation, signed language marks more categories as well as finer distinctions within them, and a number of these appear to be distinguished as well by visual parsing of rotational movement. Overall, the additional structural spatial distinctions represented in signed language appear to be ones also regularly abstracted out in visual scene parsing and, if this can be demonstrated, would show a closer connection of signed than of spoken language to visual perception.

2.2. In the Expression

The second venue, that of any single spatial expression, exhibits further respects in which signed language differs from spoken language in the apparent direction of visual scene parsing. Several of these are outlined next.

2.2.1. Iconic Clustering of Elements / Categories in the Expression

The structural elements of a scene of motion are clustered together in the classifier subsystem's representation of them in signed language more as they seem to be clustered in perception. When one views a motion event, such as a car driving bumpily along a curve past a tree, it is perceptually the same single object, the car, that exhibits all of the following characteristics: it has certain object properties as a Figure, it moves, it has a manner of motion, it describes a path of a particular contour, and it relates to other surrounding objects (the Ground) in its path of motion. The Ground object or objects are perceived as separate. Correspondingly, the classifier subsystem maintains exactly this pattern of clustering. It is the same single hand, the dominant hand, that exhibits the Figure characteristics, motion, manner, path contour, and relations to a Ground object. The other hand, the nondominant, separately represents the Ground object. All spoken languages diverge to a greater or lesser extent from this visual fidelity. Thus, consider one English counterpart of the event, the sentence *The car bumped along past the tree*. Here, the subject nominal, *the car*, separately represents the Figure object. The verb *bumped* clusters together the representation of the fact of motion and the manner of motion, while its sister constituent, the satellite *along* represents the presence of a path of translational motion. The preposition *past* represents the path conformation, while its sister constituent, the nominal *the tree*, represents the Ground. It in fact remains a mystery at this point in the investigation why all spoken languages using a preposition-like constituent to indicate path always conjoin it with the Ground nominal and basically never with the Figure nominal⁴, even though the Figure is what executes the path, and is so represented in the classifier construction of signed language.

2.2.2. Iconic Representation of Elements/Categories in the Expression

The classifier subsystem of signed language appears to be iconic with visual parsing not only in its clustering of spatial elements and categories, as just seen, but largely also in its representation of them. For example, it marks one basic category opposition, that between an entity and its activity, by using an object like the hand to represent an object, and motion of the hand to represent motion of the object. More specifically, the hand or other body part represents a structural entity (such as the Figure) - with the body part's configuration representing the identity or other properties of the entity - while movements or positionings of the body part represent properties of the entity's motion, location, or orientation. For example, the hand could be held flat to represent a planar object (e.g. a sheet of paper), or curved to represent a cup-shaped object. And, as seen, any such hand-shape as Figure could be moved along a variety of trajectories that represent particular path contours. But an alternative to this arrangement could be imagined.

⁴ As the only apparent exception, a "demoted Figure" (see Talmy 2000b, ch. 1) can acquire either of two "demotion particles" - e.g. English *with* and *of* - that mark whether the Figure's path had a "TO" or a "FROM" vector, as seen in *The fuel tank slowly filled with gas / drained of its gas*.

The handshape could represent the path of a Figure- e.g., a fist to represent a stationary location, the outstretched fingers held flat together to represent a straight line path, the fingers in a curved plane for a curved path, and the fingers alternately forward and back for a zigzag path. Meanwhile, the hand movement could represent the Figure's shape - e.g., the hand moving in a circle to represent a round Figure and in a straight line for a linear Figure. However, no such mapping of referents to their representations is found.⁵ Rather, the mapping in signed language is visually iconic: it assigns the representation of a material object in a scene to a material object in a classifier complex, for example, the hand, and the representation of the movements of that object in the scene to the movements of the hand. No such iconic correspondence is found in spoken language. Thus, while material objects are prototypically expressed by nouns in English, they are instead prototypically represented by verb roots in Atsugewi (see Talmy 2000b, ch. 1). And while path configurations are prototypically represented in Spanish by verbs, this is done by prepositions and satellites in English.

Finer forms of iconicity are also found within each branch of the broad entity-activity opposition. In fact, most of the spatial categories listed in section 3.2.5 that a classifier expression can represent are largely iconic with visual parsing. Thus, an entity's form is often represented by the form of the hand(s), its size by the compass of the hand(s), and its number by the number of digits or hands extended. And, among many other categories in the list, an entity's motive state, path contour, path length, manner of motion, and rate of motion are separately represented by corresponding behaviors of the hand(s). Spoken language, again, has only a bit of comparable iconicity. As examples, path length can be iconically represented in English by the vowel length of *way*, as in *The bird flew waay / waaaay / waaaaay up there*. Path length can also be semi-iconically represented by the number of iterations, as in *The bird flew up / up up / up up up and away*. Perhaps the number of an entity can be represented in some spoken language by a closed-class reduplication. But the great majority of spoken closed-class representations show no such iconicity.

The classifier subsystem is also iconic with visual parsing in its representation of temporal progression, specifically, that of a Figure's path trajectory. For example, when an ASL classifier expression represents "The car drove past the tree", the "past" path is shown by the Figure hand progressing from the nearer side of the Ground arm to a point beside it and on to its further side, much like the path progression one would see on viewing an actual car passing a tree. By contrast, nothing in any single closed-class path morpheme in a spoken language corresponds to such a progression. Iconicity of this sort can appear in spoken language only where a complex path is treated as a sequence of subparts, each with its own morphemic representation, as in *The vacuum cleaner is down around*

⁵ The size and shape specifiers (SASS's) in signed languages do permit movement of the hands to trace out an object's contours, but the hands cannot at the same time adopt a shape representing the object's path.

behind the clothes hamper. The classifier subsystem is further iconic with visual parsing in its extensive gradience. Many of the spatial categories listed in section 3.2.5 are largely represented in gradient form in classifier expressions. Spoken language has a bit of this, as where the vowel length of a *waaay* in English can be varied continuously. But the preponderant norm is the use of discrete spatial elements, typically incorporated into distinct morphemes. For example, insofar as they represent degree of remove, the separate forms in the series *on / next to / near / away from* represent increasing distance in quantal jumps. In the classifier subsystem, the gradient capacity of two different cognitive systems, those of visual perception and of motor control, are placed in sync, whereas the closest spoken language counterpart, the spatial portion of the closed-class subsystem, by contrast relies on the principle of discrete recombination.

2.2.3. A Narrow Time-Space Aperture in the Expression

Another way that the classifier expression in signed language may be more like visual perception is that it appears to be largely limited to representing a narrow time-space aperture. The tentative principle is that a classifier complex readily represents what would appear within a narrow scope of space and time if one were to zoom in with one's scope of perception around a Figure object, but little outside that narrowed scope. Hence, a classifier expression readily represents the Figure object as to its shape or type, any manipulator or instrument immediately adjacent to the Figure, the Figure's current state of Motion (motion or locatedness), the contour or direction of a moving Figure's path, and any Manner exhibited by the Figure as it moves. However, a classifier expression can little represent related factors occurring outside the current time, such as a prior cause or a follow-up consequence. And it can little represent even concurrent factors if they lie outside the immediate spatial ambit of the Figure, factors like the ongoing causal activity of an intentional Agent or other external instrumentality. By contrast, spoken languages can largely represent such nonlocal spatio-temporal factors within a single clause. In particular, such representation occurs readily in satellite-framed languages such as English (see Talmy 2000b, ch. 1 and 3). In representing a Motion event, this type of language regularly employs the satellite constituent (e.g. the verb particle in English) to represent the Path, and the main verb to represent a "co-event". The co-event is ancillary to the main Motion event and relates to it as its precursor, enabler, cause, manner, concomitant, consequence, or the like. Satellite-framed languages can certainly use this format to represent within-aperture situations that can also be represented by a classifier complex. Thus, English can say within a single clause - and ASL can sign within a single classifier expression - a motion event in which the Figure is moved by an adjacent manipulator, as in *I pinched some moss up off the rock* and *I pulled the 5-gallon bottle of water along the counter*. The same holds for a situation in which a moving Figure exhibits a concurrent Manner, as in *The cork bobbed past the seaweed*. But English can go on to use this same one-clause format to include the representation of co-events outside the aperture. Thus, English can here include

the representation of a prior causal event, as in *I kicked the football over the goalpost* (first I kicked the ball, then it moved over the goalpost). And it can represent a subsequent event, as in *They locked the prisoner into his cell (first they put him in, then they locked it)*. Within this same single-clause format, further, English can represent an Agent's concurrent causal activity outside any direct manipulation of the Figure, as in *I walked / ran / drove / flew the memo to the home office*. And English can represent a concurrent nonagentive cause of the Figure's motion, as in *The house burned down to the ground*. But ASL can represent none of the preceding sentences within a single classifier expression. For example, it cannot represent *I ran the memo to the home office* by, say, adopting the classifier for holding a thin flat object (thumb pressed against flat fingers) with the dominant hand and placing this atop the nondominant hand while moving forward with it as it shows alternating strokes of two downward pointed fingers to indicate running (or concurrently with any other indication of running). Instead a sequence of two expressions would likely be used, for example, first one for taking a memo, then one for a person speeding along.⁶

Comparably, in referring to a house, to represent "It burned down" one would, for example, need first to make the lexical sign for "burn up" and then (what can be treated as) a classifier expression for a structure collapsing: the hands together in an inverted "V" with the fingers interlocked and then sharply curled down. One could not represent this in a single classifier expression, say, by writhing one's fingers about as for flames as one moves them into an interlocked position and then curling them down. Though devised, these examples nevertheless show that it is physically feasible for a signed language to represent factors related to the Figure's Motion outside its immediate space-time ambit. Accordingly, the fact that signed languages, unlike spoken languages, do avoid such representations may follow from deeper structural causes, such as a greater fidelity to the characteristics of visual perception.

However apt, though, such an account leaves some facts still needing explanation. Thus, on the one hand, it makes sense that the aperture of a classifier expression is limited temporally to the present moment - this accords with our usual understanding of visual perception. But it is not clear why the aperture is also limited spatially. Visual perception is limited spatially to a narrow scope only when attention is being focused, but is otherwise able to process a wide-scoped array. Why then should classifier expressions avoid such wide spatial scope as well? Further, sign languages *can* include representation of the Ground object within a single classifier expression (typically with the nondominant hand), even where that object is not adjacent to the Figure.

⁶ The behavior here of ASL cannot be explained away on the grounds that it is simply structured like a verb-framed language, since such spoken languages typically *can* represent concurrent Manner outside a narrow aperture, in effect saying something like: "I walking / running / driving / flying carried the memo to the home office".

2.2.4. Many More Elements/Categories Representable Within a Single Expression

Although the spatiotemporal aperture that can be represented within a single classifier expression may be small compared to that in a spoken-language clause, the number of distinct factors within that aperture that can be represented is enormously greater. In fact, perhaps the most striking difference between the signed and the spoken representation of space in the expression is that the classifier system in signed language permits the representation of a vastly greater number of distinct spatial categories simultaneously and independently. A spoken language like English can separately represent only up to four or five different spatial categories with closed-class forms in a single clause. As illustrated in the sentence *The bat flew way back up into its niche in the cavern*, the verb is followed in turn by: a slot for indication of path length (with three members: “zero” for ‘neutral’, *way* for ‘relatively long’, *right* for ‘relatively short’); a slot for state of return (with two members: “zero” for ‘neutral’, *back* for ‘return’); a slot for displacement within the earth-frame (with four members: “zero” for ‘neutral’, *up* for ‘positive vertical displacement’, *down* for ‘negative vertical displacement’, *over* for ‘horizontal displacement’); a slot for geometric conformation (with many members, including *in*, *across*, *past*); and perhaps a slot for motive state and vector (with two members: “zero” for ‘neutral between location AT and motion TO’ as seen in *in / on*, and *-to* for ‘motion TO’ as seen in *into / onto*). Even a polysynthetic language like Atsugewi has closed-class slots within a single clause for only up to six spatial categories: path conformation combined with Ground type, path length, vector, deixis, state of return, and cause or manner. In contrast, by one tentative count, ASL has provision for the separate indication of thirty different spatial categories. These categories do exhibit certain cooccurrence restrictions, they differ in obligatoriness or optionality, and it is unlikely - perhaps impossible - for all thirty of them to be represented at once. Nevertheless, a sizable number of them can be represented in a single classifier expression and varied independently there. The table below lists the spatial categories that I have provisionally identified as available for concurrent independent representation. The guiding principle for positing a category has been that its elements are mutually exclusive: different elements in the same category cannot be represented together in the same classifier expression. If certain elements can be concurrently represented, they belong to different categories. Following this principle has, on the one hand, involved joining together what some sign language analyses have treated as separate factors. For example, the first category below covers equally the representation of Figure, instrument, or manipulator (handling classifier), since these three kinds of elements apparently cannot be separately represented in a single expression - one or another of them must be selected. On the other hand, the principle requires making distinctions within some categories that spoken languages treat as uniform. Thus, the single “manner” category of English must be subdivided into a category of “divertive manner” (e.g. moving along with an up-down bump) and a category of “dynamic manner” (e.g. moving along rapidly)

Spatial Structuring in Spoken and Signed Language

because these two factors can be represented concurrently and varied independently.

- A. entity properties
 - 1. identity (form or semantic category) of Figure / instrument / manipulator
 - 2. identity (form or semantic category) of Ground
 - 3. magnitude of some major entity dimension
 - 4. magnitude of a transverse dimension
 - 5. number of entities
- B. orientation properties
 - 6. an entity's rotatedness about its left-right axis ("pitch")
 - 7. an entity's rotatedness about its front-back axis ("roll")
 - 8. a. an entity's rotatedness about its top-bottom axis ("yaw")
 - b. an entity's rotatedness relative to its path of forward motion
- C. locus properties
 - 9. locus within sign space
- D. Motion properties
 - 10. motive state (moving / resting / fixed)
 - 11. internal motion (e.g. expansion/contraction, form change, wriggle, swirling)
 - 12. confined motion (e.g. straight oscillation, rotary oscillation, rotation, local wander)
 - 13. translational motion
- E. Path properties
 - 14. state of continuity (unbroken / saltatory)
 - 15. contour of path
 - 16. state of boundedness (bounded / unbounded)
 - 17. length of path
 - 18. vertical height
 - 19. horizontal distance from signer
 - 20. left-right positioning
 - 21. up-down angle ("elevation")
 - 22. left-right angle ("direction")
 - 23. transitions between motion and stationariness (e.g. normal, decelerated, abrupt as from impact)
- F. Manner properties
 - 24. divertive manner
 - 25. dynamic manner
- G. relations of Figure or Path to Ground
 - 26. path's conformation relative to Ground
 - 27. relative lengths of path before and after encounter with Ground
 - 28. Figure's path relative to the Path of a moving Ground
 - 29. Figure's proximity to Ground
 - 30. Figure's orientation relative to Ground

It seems probable that something more on the order of this number of spatial categories are concurrently analyzed out by visual processing on viewing a scene than the much smaller number present in even the most extreme spoken language patterns.

2.2.5. Elements/Categories Independently Variable in the Expression – Not in Pre-Packaged Schemas

The signed-spoken language difference just presented was mainly considered for the sheer number of distinct spatial categories that can be represented together in a single classifier expression. Now, though, we stress the corollary: their independent variability. That is, apart from certain constraints involving cooccurrence and obligatoriness in a classifier expression, a signer can generally select a category for inclusion independently of other categories, and select a member element within each category independently of other selections. For example, a classifier expression can separately include and independently vary a path's contour, length, vertical angle, horizontal angle, speed, accompanying manner, and relation to Ground object. By contrast, it was seen earlier that spoken languages largely bundle together a choice of spatial member elements within a selection of spatial categories for representation within the single complex schema that is associated with a closed-class morpheme. The lexicon of each spoken language will have available a certain number of such "prepackaged" spatial schemas, and the speaker must generally choose from among those to represent a spatial scene, even where the fit is not exact. The system of generalizing properties and processes seen in section 2.6 that apply to the set of basic schemas in the lexicon (including their plastic extension and deformation) may exist to compensate for the pre-packaging and closed stock of the schemas in any spoken language. Thus, what are largely semantic components within a single morpheme in spoken language correspond to what can be considered separate individually controllable morphemes in the signed classifier expression. Classifier expressions', apparent general lack of pre-packaging, of a fixed set of discrete basic schemas, or of a system for generalizing, extending, or deforming such basic schemas may well accord with comparable characteristics of visual parsing. That is, the visual processing of a viewed scene may tend toward the independent assessment of spatial factors without much pre-packeting of associated factors or of their plastic alteration. If shown to be the case, then signed language will once again prove to be closer to perceptual spatial structuring than spoken language is.

3. Cognitive Implications of Spoken/Signed Language Differences

The preceding comparison of the space-structuring subsystems of spoken and of signed language has shown a number of respects in which these are similar and in which they are different. It can be theorized that their common characteristics are the product of a single neural system, what can be assumed to be the core language system, while each set of distinct characteristics results from the activity of some further distinct neural system. These ideas are outlined next.

3.1. Where Signed and Spoken Language are Alike

We can first summarize and partly extend the properties above found to hold both in the closed-class subsystem of spoken language and in the classifier subsystem of signed language. Both subsystems can represent multifarious and subtly distinct spatial situations - that is, situations of objects moving or located with respect to each other in space. Both represent such spatial situations schematically and structurally. Both have basic elements that in combination make up the structural schematizations. Both group their basic elements within certain categories that themselves represent particular categories of spatial structure. Both have certain conditions on the combination of basic elements and categories into a full structural schematization. Both have conditions on the cooccurrence and sequencing of such schematizations within a larger spatial expression. Both permit semantic amplification of certain elements or parts of a schematization by open-class or lexical forms outside the schema. And in both subsystems, a spatial situation can often be conceptualized in more than one way, so that it is amenable to alternative schematizations.

3.2. Where Spoken and Signed Language Differ

First, the two language modalities have been seen to divide up into somewhat different sets of subsystems without clear one-to-one matchups. Thus, the spatial portion of the spoken language closed-class subsystem and the classifier subsystem of signed language may not be exactly corresponding counterparts, but only those parts of the two language modalities closest to each other in the representation of schematic spatial structure. Within this initial comparison, though, the classifier subsystem seems closer to the structural characteristics of visual parsing than the closed-class subsystem in the following ways: It has more basic elements, categories, and elements per category in its schematic representation of spatial structure. Its elements exhibit more iconicity with the visual in the pattern in which they are clustered in an expression, in their physical representation, in their progression through time, and in their gradient character. It can represent only a narrow temporal aperture in an expression (and only a narrow spatial aperture as well, though this difference from spoken language might not reflect visual fidelity). It can represent many more distinct elements and categories together in a single expression. It can more readily select categories and category elements independently for representation in an expression. And it avoids pre-packaged category-element combinations as well as generalizations of their range and processes for their extension or deformation.

3.3. A New Neural Model

In its strong reading, the Fodor-Chomsky model relevant here is of a complete inviolate language module in the brain, one that performs all and only the functions of language without influence from outside itself - a specifically linguistic "organ". But the evidence assembled here challenges such a model. What has here been found is that two different linguistic systems, the spoken and the signed,

both of them undeniably forms of human language, on the one hand share extensive similarities but - crucially - also exhibit substantial differences in structure and organization. A new neural model can be proposed that is sensitive to this finding. We can posit a “core” language system in the brain, more limited in scope than the Fodor-Chomsky module, that is responsible for the properties and performs the functions found to be in common across both the spoken and the signed modalities. In representing at least spatial structure, this core system would then further connect with two different outside brain systems responsible, respectively, for the properties and functions specific to each of the two language modalities. It would thus be the interaction of the core linguistic system with one of the outside systems that would underlie the full functioning of each of the two language modalities.

The particular properties and functions that the core language system would provide would include all the spoken-signed language properties in section 4.1 specific to spatial representation, though presumably in a more generic form. Thus, the core language system might have provision for: using individual unit concepts as the basis for representing broader conceptual content; grouping individual concepts into categories; associating individual concepts with overt physical representations, whether vocal or manual; combining individual concepts - and their physical representations - under certain constraints to represent a conceptual complex (i.e. the basis for morphosyntax); and establishing a subset of individual concepts as the basic schematic concepts that, in combinations, represent conceptual structure.

When in use for signed language, this core language system might then further connect with particular parts of the neural system for visual perception. I have previously called attention to the already great overlap of structural properties between spoken language and visual perception (see Talmy 2000a, ch. 2), which might speak to some neural connection already in place between the core language system and the visual system. Accordingly, the proposal here is that in the case of signed language, still further connections are brought into play, ones that might underlie the finer granularity, iconicity, gradience, and aperture limitations we have seen in signed spatial representations.

When in use for spoken language, the core language system might further connect with a putative neural system responsible for some of the characteristics present in spoken spatial representations but absent from signed ones. These could include the packeting of spatial elements into a stable closed set of patterned combinations, and a system for generalizing, extending, and deforming the packets. It is not clear why such a further system might otherwise exist but, very speculatively, one might look to see if any comparable operations hold, say, for the maintenance and modification of motor patterns.

The present proposal of a more limited core language system connecting with outlying subsystems for full language function seems more consonant with contemporary neuroscientific findings that relatively smaller neural assemblies link up in larger combinations in the subservance of any particular cognitive

function. In turn, the proposed core language system might itself be found to consist of an association and interaction of still smaller units of neural organization, many of which might in turn participate in subserving more than just language functions.

References

- Bennett, David C. 1975. *Spatial and Temporal Uses of English Prepositions: An Essay in Stratificational Semantics*. London: Longman.
- Bowerman, Melissa. 1989. Learning a Semantic System: What Role Do Cognitive Predispositions Play? In Mabel L. Rice & Richard L. Schiefelbusch (eds.) *The Teachability of Language*. Baltimore: P.H. Brookes Pub. Co.
- Clark, Herb. 1973. Space, Time, Semantics, and the Child. In Timothy E. Moore (ed.) *Cognitive Development and the Acquisition of Language*. New York: Academic Press.
- Emmorey, Karen. In Press. *Language, Cognition and the Brain: Insights from Sign Language Research*. Lawrence Erlbaum.
- Fillmore, Charles. 1968. The Case for Case. In Emmon Bach & Robert T. Harms (eds.) *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston.
- Gruber, Jeffrey S. 1965. Studies in Lexical Relations. PhD dissertation, MIT. Reprinted as part of *Lexical Structures in Syntax and Semantics*, 1976. Amsterdam: North-Holland.
- Herskovits, Annette. 1982. Space and the Prepositions in English: Regularities and Irregularities in a Complex Domain. PhD dissertation, Stanford University.
- Jackendoff, Ray. 1983. *Semantics and Cognition*. Cambridge, MA: MIT Press.
- Leech, Geoffrey. 1969. *Towards a Semantic Description of English*. New York: Longman Press, 1969.
- Liddell, Scott. Forthcoming. Sources of Meaning in ASL Classifier Predicates. In Karen Emmorey (ed.) *Perspectives on Classifier Constructions in Signed Languages* (provisional title). Likely publisher: Cambridge University Press
- Talmy, Leonard. 1983. How Language Structures Space. In Herbert L. Pick, Jr. & Linda P. Acredolo (eds.) *Spatial Orientation: Theory, Research, and Application*. New York: Plenum Press.
- Talmy, Leonard. 2000a. *Toward a Cognitive Semantics*, volume I: *Concept Structuring Systems*. Cambridge, MA: MIT Press.
- Talmy, Leonard. 2000b. *Toward a Cognitive Semantics*, volume II: *Typology and Process in Concept Structuring*. Cambridge, MA: MIT Press.
- Talmy, Leonard. In press. The representation of Spatial Structure in Spoken and Signed Language. In K. Emmorey (Ed.), *Perspectives on Classifier Constructions in Sign Languages*. Mahwah, NJ: Lawrence Erlbaum and Associates.

Leonard Talmy

Zubin, David & Soteria Svorou. 1984. Orientation and Gestalt: Conceptual Organizing Principles in the Lexicalization of Space. With S. Choi. In David Testen, Veena Mishra & Joseph Drogo. *Lexical Semantics*. Chicago: Chicago Linguistic Society.