Plans, Affordances, and Combinatory Grammar

Mark Steedman (steedman@cogsci.ed.ac.uk)
Informatics, University of Edinburgh, 2 Buccleuch Place
Edinburgh EH8 9LW, Scotland UK

Draft, May 17, 2002

Abstract

The idea that natural language grammar and planned action are related systems has been implicit in psychological theory for more than a century. However, formal theories in the two domains have have tended to look very different. This article argues that both faculties share the formal character of applicative systems based on operations corresponding to the same two combinatory operations, namely *functional composition* and *type-raising*. Viewing them in this way suggests simpler and more cognitively plausible accounts of both systems, and suggests that the language faculty evolved in the species and develops in children by a rather direct adaptation of a more primitive apparatus for planning purposive action in the world by composing affordances of objects or tools. The knowledge representation that underlies such planning is also reflected in the natural language semantics of tense, mood, and aspect, which the paper begins by arguing provides the key to understanding both systems.

1 Temporal and Causal Ontologies of Events

We are encouraged by the literature on tenses and aspects in European languages like English—not least by that large component of that literature that has appeared in the pages of the very journal whose longevity we celebrate with this anniversary number—to think of these devices as temporal in nature, to be captured by quantification over instants or intervals defined on a time-line modeled by the real numbers. However, devices like the English perfect are not primarily to do with time at all, as comparison of (1) and (2) shows:

(1) I have forgotten your name (# but I have remembered it again)

(2) Yesterday, I forgot your name (but I (have) remembered it again)

The perfect denotes a state or situation corresponding to the consequences for the subject that usually hold as a result of the core event of forgetting a name. Since these consequences include not being aware of it, and the consequences of remembering it include being aware of it, (1) is self-contradictory. Such a notion of *consequent state* is not involved in the meaning of the past tense, so the same does not hold of (2).

It is because it is hard to think of any characteristics for the state consequent upon breathing that (3a) seems infelicitous in most contexts. (3b) is similarly infelicitous because all consequences for Einstein that ensue from his visiting New York or doing anything else seem to demand his conscious existence, which we know not to hold. On the other hand, consequences of the same event for New York do not make the same demand, so (3c) is felicitous:

(3) a. #I have breathed
 b. #Einstein has visited New York
 c. New York has been visited by Einstein

For similar reasons, the Progressive is non-uniform with respect to factive entailment for core events of different aspectual type:

(4) a. Keats was writing $\models$ Keats wrote
 b. Keats was writing a sonnet $\not\models$ Keats wrote a sonnet

The observation of this "imperfective paradox" has given rise to various ontologies of events based on related syntactic tests, such as compatibility with various temporal adverbials, drawing such distinctions as that of Vendler (1967) between atelic *activities* like writing (which are compatible with modifiers like *for an hour* but not with *in an hour*) and telic *accomplishments* like writing a sonnet (which are compatible with the latter but not the former), and to to formalizations of the distinction in terms of properties like downward entailment and inertia, to deal with the fact that saying (4b) is compatible with reference to situations where one might expect a completed sonnet, whether or not it was actually produced:

(5) A proposition ϕ holding of an interval t is downward entailing if it entails that ϕ also holds of all subintervals of t down to some reasonable minimum size.

It has been commonplace since Dowty (1972, 1979) and Verkuyl (1972) to observe that such aspect is not inherent in *verbs* as Vendler may have believed, but is rather a property of *propositions*. What is more, world knowledge may contribute to the determination of the aspectual category of a proposition. For example, if we know that Keats writes a sonnet to time every Sunday afternoon, and that on a certain Sunday he wrote his sonnet in just fifteen minutes, we can say the following of that day:

(6) On 23rd December, 1916, Keats wrote in fifteen minutes.

The reason is that this specific contextual situation supports a meaning of *Keats writing* that involves an implicit object identifying a goal, the completion of the sonnet, supporting the accomplishment-demanding adverbial "in fifteen minutes".

The thoroughgoing entanglement of temporal semantics with notions of telicity and consequence makes it reasonable to consider basing formalism on representations which make goals and causation the central primitive notions, rather than time itself. This paper begins in section 2 following Steedman (1997) by seeking such representations among those used in artificial intelligence for the construction of plans of action, rather than among standard temporal logics, although many affinities between logicist and computational approaches will be apparent. It is from these systems that parallels with grammar discussed in sections 3 and 4 will emerge, via the further involvement of the Gibsonian notion of "affordance" of actions by objects in the world. The concluding section 5 briefly examines the implication that the language faculty is built upon a more primitive planning faculty, and some evidence from neurology and neuroanatomy for the reality of this parallel.

2 The Linear Dynamic Event Calculus

The Linear Dynamic Event Calculus (LDEC, cf. Steedman 1995, 1997) combines the insights of the Event Calculus of Kowalski and Sergot (1986), itself a descendant of the Situation Calculus of McCarthy and Hayes (1969) and the planner of Fikes and Nilsson (1971), with the Dynamic Logics that were developed by Harel (1984) and others for reasoning about computer programs.[1] Dynamic logics are a form of modal logic in which the $\Box$ and $\Diamond$ modalities are relativized to particular events. For example, if a (possibly nondeterministic) program or command α computes a function F over the integers, then we may write the following:

(7) $n \geq 0 \Rightarrow [\alpha](y = F(n))$

(8) $n \geq 0 \Rightarrow \langle\alpha\rangle(y = F(n))$

The intended meaning of the first of these is "in any situation in which $n \geq 0$, after every execution of α that terminates, $y = F(n)$". That of the second is (dually) that "in any situation in which $n \geq 0$, there is an execution of α that terminates with $y = F(n)$".

We can think of these modalities as defining a logic whose models are Kripke diagrams in which accessibility between possible worlds is represented by individual events, as in figure 1. Such events can be defined as mappings between situations or partially specified possible worlds,

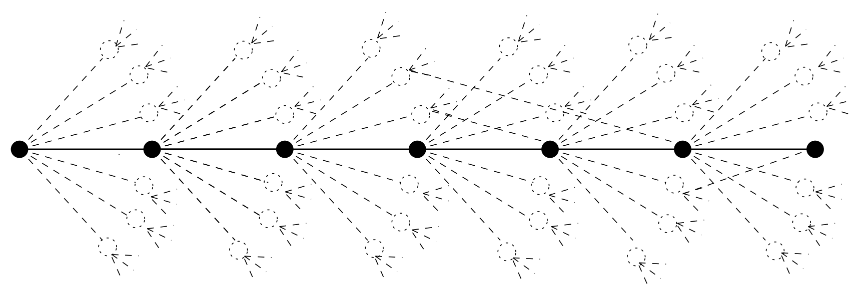


Figure 1: Kripke Model of Causal Accessibility Relation

defined in terms of conditions on the antecedent which must hold for them to apply (such as that $n \geq 0$ in (7)), and consequences (such as that $y = F(n)$) that hold in the consequent.

The particular dynamic logic that we are dealing with here is one that includes the following dynamic axiom, which says that the operator ; is *sequence*, an operation related to functional composition over events, viewed as functions from situations to situations:

(9) $[\alpha][\beta]P \Rightarrow [\alpha;\beta]P$

The situation calculus and its many variants can be seen as "reified" versions of this dynamic logic.

Composition is one of the most primitive *combinators*, or operations combining functions, which Curry and Feys (1958) call **B**. It can be defined by the following equivalence with a lambda term:

[1]LDEC is closely related to a number of other descendents of the situation and event calculi, including Gelfond and Lifschitz (1993), Sandewall (1994), Shanahan (1997) and Thielscher (1998, 1999).

(10) $\mathbf{B}\alpha\beta \equiv \lambda s.\alpha(\beta s)$

Accordingly, the above sequence $\alpha;\beta$ could be written in this notation as $\mathbf{B}\alpha\beta$

The situation calculi are heir to a problem known in the AI literature as the Frame Problem. This problem arises because the way that we structure our knowledge of change in the world is in terms of event-types characterized (mostly) by *localized* effects. For example, if I eat a hamburger, the effects of the action are confined to the hamburger and myself: most of the myriad other facts which characterize the situation such as the color of the walls and the position of every other object in the room are unaffected. (This is not a logical necessity in an event representation: it just seems to be the way we think about the world, as a consequence of our physical being in it.) The frame problem arises in two forms: a "representational" form and an "inferential" or computational form.

The representational frame problem is simply that it is cumbersome and tedious to have to explicitly state the fact that eating hamburgers leaves unchanged the color of the walls, the position of each table and each chair, etc. via "frame axioms" like the following:[2]

(11) $color(wall,x) \Rightarrow [eat(hamburger)]color(wall,x)$

Many of these calculi elegantly solve the representational frame problem with a device that seems to have first been proposed by Kowalski (1979), namely frame axioms for each action-type saying such things as that all facts about the world which do not involve the presence of the intact hamburger or my being hungry remain as they were when I eat a hamburger, thus:[3]

(12) $p \wedge (p \neq here(hamburger)) \Rightarrow [eat(hamburger)]p$

This solution is helpful in that it keeps rules describing eating hamburgers as simple as rules like (7), and avoids explicit representation of a lot of facts most of which will be entirely irrelevant to any given proof or any given plan. However, if we ever need to know what color the walls are after a sequence of hamburger eating actions, then it will still be necessary to explicitly chain through the relevant instances of the frame axioms to establish that the color of the walls is still as it always was.

This second problem is the *inferential* form of the frame problem. It is interesting that the very first computationally practicable event representation, which is generally referred to as the STRIPS representation after its use in the program of that name by Fikes and Nilsson (1971), solves both forms of the frame problem. It does so by representing the state of the world as the equivalent of a Prolog database of facts like $here(hamburger99)$, $hungry(me)$, and $pink(walls)$. Events like eating something were represented by *preconditions*, and by sets of *additions* and *deletions* which explicitly added and deleted facts to and from the database, as in the following operator capturing the fact that eating a hamburger makes you thirsty and makes the hamburger no longer available:

[2]We follow standard logic programming conventions under which variables in the consequent to the right of the implication arrow are implicitly universally quantified and any other variables in the antecedent are implicitly existentially quantified.

[3]This frame axiom, which anticipates the idea of "explanation closure" (Schubert 1994), follows the logic programming convention of negation by failure. That is, the rule does not say that the door is neither open nor shut after you push it, but simply that a proof of either property must be sought elsewhere—say via rules (14).

(13) OPERATOR: $\lambda x.eat(x)$
PRECONDITIONS: $hamburger(x)$
$here(x)$
$hungry$
DELETIONS: $here(x)$
ADDITIONS: $thirsty$

As a result, if STRIPS needed to know the color of the walls or where I am at the end of a sequence of hamburger-eating actions, it could do it by database lookup, rather than an expensive chain of inferences.

STRIPS avoids the inferential frame problem by modeling the inertia of the real world by the inertia of database update. While STRIPS itself is expressed in procedural terms, a number of recent logicist theories of action representation, notably Bibel et al. (1989), Martí-Oliet and Meseguer (1999), and Thielscher (1999), have proposed building the same insight into logical representations of events.

To do this, we need a new form of logical implication, distinct from the standard or intuitionistic $\Rightarrow$ we have used up till now. We will follow Girard (1987, 1995), Bibel et al. (1989), Japaridze (1998), and others in using *linear* logical implication $\multimap$ rather than intuitionistic implication $\Rightarrow$ in those rules that change the value of fluents.

For example, we can represent events involving doors in a world (simplified for purposes of exposition) in which there are two places *out* and *in* separated by a door which may be *open* or *shut*, as follows:[4]

(14) a. $shut(x) \multimap [push(x)]open(x)$
b. $open(x) \multimap [push(x)]shut(x)$

(15) a. $in \multimap [go\text{-}through(x)]out$
b. $out \multimap [go\text{-}through(x)]in$

Linear implication has the effect of building into the representation the update effects of actions—that once you apply the rule, the proposition in question is "used up", and cannot take part in any further proofs, while a new fact is added. The formulae therefore say that if something is shut and you push it, it becomes open (and vice versa), and that if you are in and you go through something then you become out (and vice versa).

To interpret linear implication as it is used here in terms of proof theory and proof search, we need to think of possible worlds in the Kripke diagram in figure 1 as states of a single updatable STRIPS database of facts. Rules like (14) and (15) can then be interpreted as (partial) functions over the states in the model that map states to other states by removing facts and adding other facts. Linear implication and the dynamic box operator are here essentially used as a single state-changing operator: you can't have one without the other.

We will pass over here the precise way that this procedure can be captured within the logic, so as to yield desirable properties such as a cut elimination theorem and the related subformula property, which are essential if a logic is to be usable as a logic programming language and to support practicable proof-search. See Thielscher (1999), Martí-Oliet and Meseguer (1999), and

[4]Since in the real world doors don't always open when you push them, it might seem more appropriate to use diamond modalities such as $\langle push(x)\rangle$. However, this gets us into irrelevant and in general insoluble "qualification problems" so I present a simplified deterministic version here, in which box can be read as *default* necessity, meaning "usually".

Pym (2001) for computationally practical proof theories and semantics for systems of a similar kind.

The effect of such systems can be exemplified as follows. If the initial situation is that you are in and the door is shut:

(16) $in \wedge door(d) \wedge shut(d)$

—then the linear rules (14) mean that an attempt to prove the proposition in (17) concerning the state of the door in the situation that results from pushing the door will fail because rule (14a) has removed the fact in question from the database that results from the action $push(d)$:[5]

(17) $[push(d)]shut(d)$

On the other hand, attempts to prove the following will all succeed, since they are all facts in the database that results from the action $push(d)$ in the initial situation (16):

(18) a. $[push(d)]open(d)$
b. $[push(d)]door(d)$
c. $[push(d)]in$

The advantage of interpreting linear implication in this way is that it builds the STRIPS treatment of the frame problem (Fikes and Nilsson 1971) into the proof theory, and entirely avoids the need for inferentially cumbersome reified frame axioms like (12).

Using linear implication (or the equivalent rewriting logic devices or state update axioms of Thielscher (1999) and Martí-Oliet and Meseguer (1999)) for STRIPS-like rules makes such frame axioms unnecessary. Instead, they are theorems concerning the linear logic representation. The further implications of the theory for extended forms of the frame problem considered by Hanks and McDermott (1986), Sandewall (1994) and Shanahan (1997) are discussed in Steedman 1997, 2000a.

Even in this extremely simplified world, we need a little more apparatus to represent our knowledge about doors in a way which will allow us to make plans involving them. We also need to state preconditions on the actions of pushing and going through. Here ordinary non-linear intuitionistic implication is appropriate:[6]

(19) a. $door(x) \wedge open(x)$
$\Rightarrow possible(go\text{-}through(x))$
b. $door(x) \Rightarrow possible(push(x))$

These rules say (oversimplifying wildly) that if a thing is a door and is open then it's possible to go through it, and that if a thing is a door then it's possible to push it.

We also need to define the transitive property of the possibility relation, as follows, using the definition (9) of event sequence composition:

(20) $possible(\alpha) \wedge [\alpha]possible(\beta) \Rightarrow possible(\alpha;\beta)$

[5]We follow the logic programming convention of negation by failure, according to which a proposition is treated as false if it cannot be positively proved to be true.

[6]The version of linear logic mixing linear and standard implication is is a hybrid or "multi-modal" logic of the kind investigated by Moortgat (1997) and others. This particular logic is closely related to a non-dynamic version investigated by O'Hearn and Pym (1999) and Pym (1999) under the name of "Bunched Implication Logic". Pym (2001) gives an extensive treatment of its semantics and proof theory, including a cut elimination theorem, and Armelín and Pym (2001) discuss its use for practical logic programming.

This says that any situation in which it is possible to α, and in which actually doing α gets you to a situation where it is possible to β, is a situation in which it is possible to α *then* β.

If we regard actions as functions from situations to situations, then this rule defines *function composition* as the basic plan-building operator of the system. Composition is one of the simplest of a small collection of *combinators* or operations for combining functions to yield new functions, which Curry and Feys (1958) used to define the foundations of the λ-calculus and other applicative systems in which new concepts can be defined in terms of old. Since the knowledge representation that underlies human cognition and human language could hardly be anything *other* than an applicative system of some kind, we should not be surprised to see it turn up as one of the basic operations of planning systems.

This fragment gives us a simple planner in which starting from the world (21) in which I am *in*, and the door is *shut* and stating the goal (22) meaning "find a possible series of actions that will get me *out*," can, given a suitable search control (see Cervesato et al. 1996), be made to automatically deliver a constructive proof that one such plan is (23):

(21) $in \wedge door(d) \wedge shut(d)$

(22) $possible(\alpha) \wedge [\alpha]out$

(23) $\alpha = push(d); go\text{-}through(d).$

One way to produce this proof, which is suggested as an exercise, is via *backward-chaining* from the goal (22) on the consequents of rules (19) using the transitivity rule (20). The situation that results from executing this plan in the start situation (16) is one in which the following conjunction of facts is directly represented by the database:

(24) $out \wedge door(d) \wedge open(d)$

Slightly more interestingly, we can represent the knowledge required to solve the "monkey and bananas" problem (simplifying as usual) as follows.

First, we represent the fact that the consequence of grabbing something is that you have it, and that (somewhat grotesquely, in order to shorten the proof by avoiding some trivial arithmetic) if you are at a position six feet above where you are now, you can grab the bananas:

(25) a. $[grab(x)]have(x)$
b. $at((here + 3) + 3) \Rightarrow possible(grab(bananas))$

Next, we define the conditions and side-effects of getting on a box. If something is a box you can climb on it:

(26) $box(b) \Rightarrow possible(climb\text{-}on(b))$

—and if you are at a place and you climb on a box you are at a place that is higher by 3ft:

(27) $at(p) \multimap [climb\text{-}on(b)]at(p+3)$

We do the same for the action of putting boxes on things: if two things have nothing on top of them and are not the same thing you can put one on the other:

(28) $clear(x) \wedge clear(y) \wedge (x \neq y) \Rightarrow possible(puton(x,y))$

—and if x is on something and you put it on something else then that something becomes clear and x is on that something else:

(29) a. $on(x,z) \multimap [puton(x,y)]clear(z)$
b. $clear(y) \multimap [puton(x,y)]on(x,y)$

(It is worth noting that the use of a hybrid logic in which fluents like *clear* can be antecedents to both nonlinear qualification rules like (28) and linear ramification rules like (29) means that we avoid reintroducing the frame problem by having to explicitly state that $clear(x)$ in the consequent of (29b), an outcome which the pure linear systems of Martí-Oliet and Meseguer (1999) and Hölldobler and Sneeberger (1990) have difficulty getting around—see Martí-Oliet and Meseguer 1999, 44-46.)

If the initial state of the world is as follows:

(30) $at(here) \wedge box(b1) \wedge box(b2) \wedge clear(b1) \wedge clear(b2)$

—then the goal (31a) gives rise to (31b) as one possible plan:

(31) a. $possible(\alpha) \wedge [\alpha]have(bananas)$
b. $\alpha = [puton(b1, here); climb\text{-}on(b1);$
$puton(b2, b1); climb\text{-}on(b2); grab(bananas)]$

However, we have said nothing yet about the problem of *Search* implicit in searching for and identifying such plans.

3 Planning, Temporality, and Affordance in LDEC

Although the example is simplified for purposes of exposition, it provides the basis for a quite general calculus of events. (See Shanahan (1997), Thielscher 1999, and Steedman (1997, 2000b) for related proposals including discussions of ramification, qualification, delayed action, simultaneity, nondeterminism and other standard problems that such representations have to deal with.)

In particular, it provides the basis for the primary data structure underlying the event representation proposed by Moens (1987) and Moens and Steedman (1988) (see Figure 2), who explained the coercion of *aktionsarten* or aspectual types and the semantics of temporal adverbial modifiers influentially described by Vendler (1967), including the aspecual "paradoxes" illustrated in examples (1) and (4) in terms of a knowledge representation associating events with preparatory activities and consequent states, and a system of lexically controlled transitions among these associates. (The approach of Narayanan (1997, 1999) using Petri nets is closely

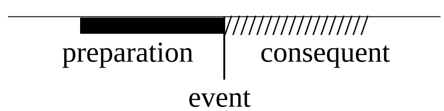


Figure 2: The event nucleus (adapted from Moens and Steedman 1988)

related.)

Durative events can then be represented in terms of inceptive and terminative events with progressive states as their respectively consequent and preparatory states.

In fact the representation of actions and events in terms of an association of preconditions and consequences with the core event is a very generally applicable one. If the precondition is a conditional stimulus such as a light, and the consequence is a reward, such as food, while the action concerned is pecking or pressing a bar, then it can be considered as a representation of an *operant* in the cognitive sense of Rescorla and Wagner (1972), itself a notion closely related to that of an *affordance* (Gibson 1966).

The notion of an affordance has been used in two rather different ways. In its most basic sense of an invariant supporting perception, it has been extremely helpful in directing attention to non-obvious properties of the sensory array relevant to visual and haptic perception, and motor control (Lee 1980; Turvey 1990). In its more general sense of an interaction with the world that a perceived object mediates (Gibson 1979) it has proved equally attractive to a wide range of theoretical positions that have emphasized the fundamental role of the notion of *action* in human cognition (Norman 1988, 1999). This is the sense in which a door "affords" egress and ingress, a knife affords cutting and scraping, and the like. The attraction of this notion is that it seems to offer a way in which perceptual learning can be linked to the goals and actions upon the environment of the learner, an idea that has been followed up by E. Gibson and Spelke (1983), among others. However, its influence in these domains has been limited by two difficulties.

One has been the controversial idea of "direct perception"—the claim that affordances of this generalized kind are as directly signalled by properties of the sensory array as are affordances in the narrower sense, such as time-to-impact with the surface of the sea for a diving gannet, an affordance which has been plausibly identified by Lee (1980) with specific parameters of the optic flow field.[7] It is certainly hard to believe that the perception that a postbox "affords letter-mailing to a letter-writing human in a community with a postal system" (Gibson 1979, p.139, citing Gibson 1950) is "direct" in anything like the same sense, although recognition of postboxes, like that of everything else, is undoubtedly mediated *in part* by such Gibsonian invariants of the optic array as relative spatial frequency spectra.[8] In fact, it seems unlikely that Gibson himself intended any such strong claim. I shall ignore the perceptual aspect of affordances in the theory developed below.

A more serious obstacle to the exploitation of the idea of affordances in the more general sense has stemmed from the very fact that many such affordances are actions or events. A formal theory of events in their relation to objects that is applicable to human cognition and/or natural language semantics has been lacking. It is this lack that the present paper seeks to address.[9]

LDEC offers a basis for a formalization of the relation between objects and their affordances, of the kind that we need in order to talk about perceptual and cognitive learning in non-linguistic animals and prelinguistic children. For example, the facts in (14) and (15) strike me as a fair approximation to what my cat knows about the affordances of doors. Of course, the representa-

[7]This seems to be the reason why Lakoff (1987), who might otherwise have been expected to be sympathetic to the idea, rejected affordances as a basis for human conceptual categorization. I shall suggest below that this rejection was over-hasty.

[8]This scepticism is reinforced by the experience of moving to the United States, when it took some years of perceptual learning and painful cognitive effort at all levels before the affordances of US Mail postboxes became perceptually available in the way that those of the functionally identical British ones were.

[9]To say this is not to deny the existence and potential relevance of Tense Logic, Temporal Logic, and other calculi over situations and events, or the related AI systems for planning and reasoning about action, upon all of which the present approach builds, but rather to say that their application to human cognition in general and generalized affordances in particular is far from straightforward.

tion says nothing about the invariants that afford the perception of doors in the first place, their relation to bodily properties like the size of my cat's head, her grounding or motor embedding of the actions of pushing and going through, and so on. It represents *that which is* perceived or learned. Nevertheless, the representation could be used to explain the transition she seemed to go through through from a stage where doors afforded her (15) (going through for purposes of egress and ingress) but not (14) (pushing to open and close), homing in via a set of superstitious and rapidly extinguishing spurious affordances to a correct affordance (14) supporting the motor plan (23) and possibly its internalization as yet another affordance of doors. The representation also suggests a basis for experimental investigation of the precise details of my cat's representation of the affordances of doors. (For example, do they afford her the ingress and egress of other cats? Does her cat-door anomalously afford her *my* egress?)

Many of these experiments have already been done for us—most notably, by Köhler (1925), in his investigations of tool use and planning in Chimpanzees, including their ability to produce plans like (31b), among other even more exuberant tool-based solutions to problems with bananas (see Figure 3).

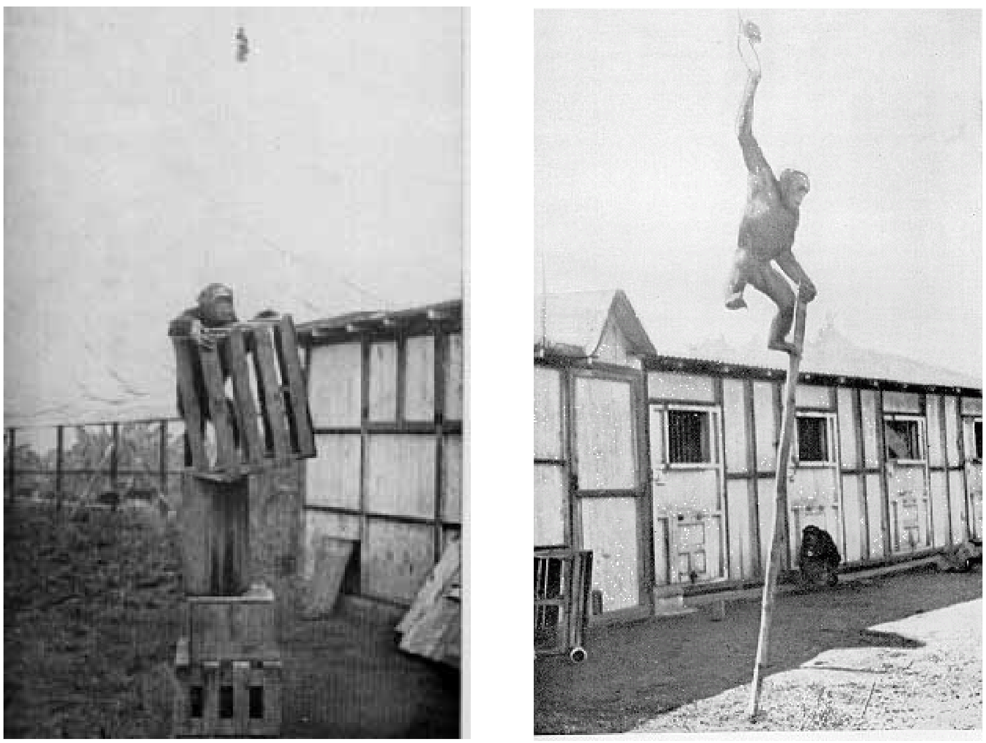


Figure 3: Chimpanzee plans (from Köhler 1925, reprinted by permission of the publisher)

One of Köhler's most thought-provoking observations was the following. A chimpanzee which was perfectly capable of consistently using a tool such as a stick to reach otherwise unattainable objects—one to whom sticks afforded reaching—was unable to enact such a plan unless the stick was actually present in the problem situation. Mere availability of a stick in an adjoining room—even one which the ape had recently explored—was not enough to trigger the relevant knowledge and cause the ape to fetch the stick.

This observation suggests that for non-linguistic animals, even those closest to us in evolutionary terms, access to the affordances of objects is tied to immediate perception of the objects themselves, as Gibson believed. For an animal, this is quite a good way of running your planner. If you don't have much control over your physical environment, it is probably better to look at those plans the situation affords, rather than backward chaining to conditions that there may be no way for you to satisfy, say because of the time of year. This in turn suggests, uncontroversially, that affordances like egress are indexed in such animals by object-concepts like *door*, rather than by end-states like being *out*, and that planning proceeds *reactively* (Schoppers 1987) by forward chaining from what is the case, rather than backward chaining from the goal.

We can represent such indexing by first defining actions like *pushing* and *going through* as functions like the following derived from (14) and (15):[10]

(32) a. $push(x) = \lambda x.\left\{ \begin{array}{l} shut(x) \multimap open(x) \\ open(x) \multimap shut(x) \end{array} \right\}$

b. $go\text{-}through(x) = \lambda x.\left\{ \begin{array}{l} in \multimap out \\ out \multimap in \end{array} \right\}$

(Here the linear implication symbol $\multimap$ is overloaded to signify linear mapping of state to state accompanied by deletion and addition of facts. Implication is so closely related to functional mapping, and the functions in question are so closely related to the state update or rewrite axioms of the proof theory that this overloading seems unlikely to cause confusion.)

The set of such functions $affordances(door)$ constitutes the affordances of doors:

(33) $affordances(door) = \left\{ \begin{array}{l} push \\ go\text{-}through \end{array} \right\}$

The Gibsonian affordance-based door-schema can then in turn be defined as a function mapping doors into (second-order) functions from their affordances like pushing and going-through to their results:

(34) $door' = \lambda x_{door}.\lambda p_{affordances(door)}.px$

The operation of turning an object of a given type into a function over those functions that apply to objects of that type is another primitive combinator called **T** or *type-raising*. As in the case of composition (10), the effect of this combinator can be defined by equivalence to the corresponding λ-term:[11]

(35) $\mathbf{T}x \equiv \lambda p.px$

[10] A first sketch for this analysis appears in Steedman 2002.

[11] Combinators like (10) and (35) were used by Curry and Feys (1958) to define the foundations of all applicative systems including the λ-calculi. (They used the name $\mathbf{C}_*$ for the combinator **T**.) Combinatory systems that include both composition and type-raising are quite expressive—see Smullyan (1985, 1994) for discussion. They have the character of calculi for rebracketing and permuting terms in expressions, and it is this property that motivates the linguistic use of some related operations in section 4 below. Such calculi are closely related to linear logic itself—see van Benthem (1991) for discussion. The combinator **T** is also related to the notions of object-orientation and continuation-passing in programming language theory—see Barker (2001) for discussion of the linguistic and cognitive significance of the latter.

Accordingly, (34) can be rewritten:

(36) $door' = \lambda x_{door}.\mathbf{T}x$

Such a concept of doors is useful for reactive planning, and one can add more affordances to *affordances*(*door*) as one's experience increases. It seems quite likely that this is close to the way cats or at least chimpanzees conceptualize doors.

It is interesting in this connection that "The Sims"—a currently wildly popular interactive computer game involving multiple semi-autonomous animated humanoid agents or characters in a domestic and social setting—owes much of the appearance of purposeful autonomous action in its characters—such as it is—to just such a representation. That is, proximity to doors makes available to the agent actions of opening and closing, that of refrigerators those of getting a beer, showers those of removing clothes and taking a shower, and so on. Coupled with a simple representations of the agent's internal state and a limited forward-chaining planning ability, this is enough to convey an impression of intentional action that many evidently find compelling.

However, such representations are in human terms somewhat stultifying, in that they restrict one's door concept to previously encountered events involving doors that one has somehow stumbled across. One would like to have the advantages in terms of efficiency of planning that thinking of objects in terms of their affordances allows, while also being able envisage novel uses for doors—for example, using one as a table, or as a raft—when circumstances demand it. In other words, one would like to be able to generalize (34) over a wider range of affordances, such as the affordances of flat movable objects, or of other things that you can push and/or go through.

At this point we encounter some very well-known deep and intractable problems in the study of cognition. We know very little about the way that people conceptualize objects so as to permit fruitful generalization. There are reasons to think our ability to generalize very far beyond directly experienced affordances is quite limited. (For example, people find considerable difficulty in solving those irritating conundrums which require you to see a pair of pliers as the weight for a plumbline, or the box that thumbtacks are packaged in as a bracket that can be pinned to a wall with the thumbtacks in order to support a candle.) It seems likely that the basis for such limited generalization is partly perceptual, and partly embedded in our modes of interaction with objects, as Gibson insisted.

One of the few sources of information about the natural classifications of objects that permit limited generalization comes from linguistics. For example, many North American Indian languages, such as the Athabascan group that includes Navaho, are comparatively poorly-off for nouns. Many nouns for artefacts are morphological derivatives of verbs. For example, "towel" is *bee 'ádít'oodí*, glossed as "one wipes oneself with it", and "towelrack" is *bee 'ádít'oodí bąąh dah náhidiiltsos*—roughly "one wipes oneself with it is repeatedly hung on it". Since the verbs themselves are agglutinations of comparatively abstract action elements and are marked for the physical nature of their complements by pronominal or agreement-like morphemes, these languages have the appearance of having rather directly conventionalized a paradigm affordance as the lexical item for the artefact.[12]

[12] I am grateful to Joyce McDonough for advice on this topic. See Young and Morgan 1987, 7 for details. It is important to be clear that while the morphological processes that produce such lexical items are completely productive, the nominalizations themselves are highly conventionalized and behave in many respects exactly like our own more arbitrary lexical items. This observation suggests that Navaho morphological composition is, like that of English, largely

One might also view morphological *case* (and possibly other morphological agreement markers, such as the extended *gender* systems of some Bantu languages) as reflecting nominal type-raising over verbal functions that take classes of objects such as those that can participate in the role(s) defined by nominative, agentive, accusative, ergative etc.

The next section argues that *all* languages, whether or not they bear explicit morphological indicators of these kinds, involve both type-raising of nominal categories and composition of verbal ones as assumed in the Combinatory Categorial Grammar (CCG) framework. This suggests both that grammatical phenomena can be used to investigate the fine detail of cognitive representations, and that the fundamental operations of linguistic syntax are drawn from more general and more cognitively primitive operations.

4 Combinatory Grammars.

Combinatory Categorial Grammar (CCG, Ades and Steedman 1982, Steedman 1990, Steedman 2000a), like other varieties of Categorial Grammar reviewed in Wood (1993) and exemplified in the bibliography below, is a theory in which all linguistic elements are categorized or typed as either functions or basic types, and in which syntactic derivation is achieved by syntactic rules corresponding to directionally restricted versions of a small number of combinators prominently including composition **B** and **T**. Thus it is a theory that makes language look as if it has been attached to a pre-existing system for planning action in the world, and thereby seem less innately distinct and evolutionarily singular as a cognitive faculty than is usually assumed. It is perhaps worth digressing briefly to recall the specific form in which these operations show up in language, according to this theory.[13]

In CCG elements like verbs are associated with a syntactic "category" which identifies them as *functions*, and specifies the type and directionality of their arguments and the type of their result. For example, a transitive verb is a function from (object) NPs on the right into predicates—that is, into functions from (subject) NPs on the left into S:[14]

(37) likes := $(S\backslash NP)/NP$

(38) *Forward Application:* (>)
$$X/Y \quad Y \quad \Rightarrow \quad X$$

(39) *Backward Application:* (<)
$$Y \quad X\backslash Y \quad \Rightarrow \quad X$$

These rules have the form of very general binary PS rule schemata. In fact, pure categorial grammar is just context-free grammar written in the accepting, rather than the producing, direction, with a consequent transfer of the major burden of specifying particular grammars from the PS

"off-line" and not an active component of sentence processing. They also imply that any claim that these and similar languages draw no clear distinction between nouns and verbs is of dubious significance. (Here aphasia data of the kind discussed in section 5 might be very revealing.) Of course Navaho speakers' ability to reason about other uses of towels is no more limited by the literal interpretation of the conventionalized form than that of English speakers.

[13] Much of this section summarizes ideas that were originally presented in the pages of *Linguistics & Philosophy*, and can be taken as read or skimmed by those who have read it attentively over past years.

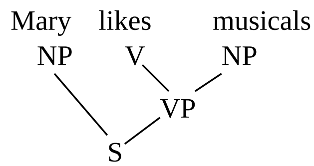
[14] We here use the "result leftmost" notation in which a rightward-combining functor over a domain β into a range α are written α/β, while the corresponding leftward-combining functor is written $\alpha\backslash\beta$. (α and β may themselves be function categories.) There is an alternative "result on top" notation due to Lambek (1958), according to which the latter category is written $\beta\backslash\alpha$.

rules to the lexicon. While it is now convenient to write derivations as in a, below, they are equivalent to conventional phrase structure derivations b:

(40) a. Mary likes musicals

$$\frac{\dfrac{\text{Mary}}{NP} \quad \dfrac{\dfrac{\text{likes}}{(S\backslash NP)/NP} \quad \dfrac{\text{musicals}}{NP}}{S\backslash NP}>}{S}<$$

b.



It is important to note that such tree-structures are simply a representation of the process of derivation. They are not structures that need to be built by a processor, nor do they provide the input to any rules of grammar.

Such categories can be regarded as encoding the semantic type of their translation, and this translation can be made explicit in the following expanded notation, which associates a translation with the entire syntactic category, via the colon operator, which is assumed to have lower precedence than the categorial slash operators. (Agreement features are also included in the syntactic category, represented as subscripts, much as in Bach 1983. The feature 3*s* is "underspecified" for gender and can combine with the more specified *3sm* by a standard unification mechanism that we will pass over here – see Shieber 1986.)[15]

(41) likes := $(S\backslash NP_{3s})/NP : like'$

We must also expand the rules of functional application in the same way:

(42) *Forward Application:* $(>)$
$$X/Y : f \quad Y : a \quad \Rightarrow \quad X : fa$$

(43) *Backward Application:* $(<)$
$$Y : a \quad X\backslash Y : f \quad \Rightarrow \quad X : fa$$

They yield derivations like the following:

(44) Mary likes musicals

$$\frac{\dfrac{\text{Mary}}{NP_{3sm} : mary'} \quad \dfrac{\dfrac{\text{likes}}{(S\backslash NP_{3s})/NP : like'} \quad \dfrac{\text{musicals}}{NP : musicals'}}{S\backslash NP_{3s} : like'musicals'}>}{S : like'musicals'mary}<$$

The derivation yields an S with a compositional interpretation, equivalent under a convention of left associativity to $(like'musicals')mary'$.

Coordination might be included in CG via the following rule, allowing constituents of like type to conjoin to yield a single constituent of the same type:[16]

[15] This notation follows Steedman 1987. Another notation, used in Steedman 1990, associates a unifiable logical form with each primitive category, so that the same transitive verb appears as follows:

(i) likes := $(S : like'y\,x\backslash NP_{3s} : x)/NP : y$

The advantage is that the predicate-argument structure is built directly by the unification, and that the combination rules need no further modification. Otherwise the choice is largely a matter of notational convenience.

[16] The semantics of this rule, or rather rule schema, is somewhat complex, and is omitted here. The rule is also simplified syntactically in several respects for the present purpose.

(45) *Coordination:* $(<\&>)$

$$X \quad conj \quad X \quad \Rightarrow \quad X$$

(46)
$$\begin{array}{ccccc}
\text{I} & \text{loathe} & \text{and} & \text{detest} & \text{opera} \\
\overline{NP} & \overline{(S\backslash NP)/NP} & \overline{CONJ} & \overline{(S\backslash NP)/NP} & \overline{NP} \\
 & \multicolumn{3}{c}{\underline{\hphantom{(S\backslash NP)/NP\ CONJ\ (S\backslash NP)/NP}}_{<\Phi>}} & \\
 & \multicolumn{3}{c}{(S\backslash NP)/NP} & \\
 & \multicolumn{4}{c}{\underline{\hphantom{(S\backslash NP)/NP\ CONJ\ (S\backslash NP)/NP\ NP}}_{>}} \\
 & \multicolumn{4}{c}{S\backslash NP} \\
\multicolumn{5}{c}{\underline{\hphantom{NP\ (S\backslash NP)/NP\ CONJ\ (S\backslash NP)/NP\ NP}}_{<}} \\
\multicolumn{5}{c}{S}
\end{array}$$

In order to allow coordination of contiguous strings that do not constitute constituents, CCG allows certain further operations on functions related to Curry's combinators (Curry and Feys 1958). For example, functions may nondeterministically *compose*, as well as apply, under the following rule:

(47) *Forward Composition:* $(>\mathbf{B})$

$$X/Y \quad Y/Z \quad \Rightarrow \quad X/Z$$

The most important single property of combinatory rules like this is that their semantics is completely determined under the following principle:[17]

(48) *The Principle of Combinatory Transparency:* The semantic interpretation of the category resulting from a combinatory rule is uniquely determined by the interpretation of the slash in a category as a mapping between two sets.

In the above case, the category X/Y is a mapping of Y into X and the category Y/Z is that of a mapping from Z into Y. Since the two occurrences of Y identify the *same* set, the result category X/Z is that mapping from Z to X which constitutes the composition of the input functions. It follows that the only semantics that we are allowed to assign, when the rule is written in full, is as follows:

(49) *Forward Composition:* $(>\mathbf{B})$

$$X/Y : f \quad Y/Z : g \quad \Rightarrow \quad X/Z : \mathbf{B}fg$$

—or equivalently

(50) *Forward Composition:* $(>\mathbf{B})$

$$X/Y : f \quad Y/Z : g \quad \Rightarrow \quad X/Z : \lambda x.f(gx)$$

No other interpretation is allowed.[18]

The operation of this rule in derivations is indicated by an underline indexed $>\mathbf{B}$ because it is one of the two rightward-looking composition rules. Its effect can be seen in the derivation of sentences like *I requested, and would prefer, musicals*, which crucially involves the composition of two verbs to yield a composite of the same category as a transitive verb (the rest of the derivation is given in the simpler notation). It is important to observe that composition also yields an appropriate interpretation for the composite verb *would prefer*, as $\lambda x.\lambda y.will'(prefer'\ x)\ y$, an object which if applied to an object *musicals* and a subject *I* yields

[17] This principle is stated differently in Steedman (2000a) but is in fact identical.

[18] It is worth noticing that this principle would follow automatically if we were using the alternative unification-based notation discussed in note 15 and the composition rule as as it is given in (47).

the proposition $will'(prefer'\ musicals')\ me'$. The coordination will therefore yield an appropriate semantic interpretation.[19]

(51)
$$\begin{array}{cccccc}
\text{I} & \text{requested} & \text{and} & \text{would} & \text{prefer} & \text{musicals}\\
\overline{NP} & \overline{(S\backslash NP)/NP} & \overline{CONJ} & \overline{(S\backslash NP)/VP : will'} & \overline{VP/NP : prefer'} & \overline{NP}\\
 & & & \multicolumn{2}{c}{\overline{(S\backslash NP)/NP : \lambda x.\lambda y.will'(prefer'x)y}^{\;>\mathbf{B}}} & \\
 & \multicolumn{4}{c}{\overline{(S\backslash NP)/NP}^{\;<\Phi>}} & \\
 & & & \multicolumn{3}{c}{\overline{S\backslash NP}^{\;>}}\\
\multicolumn{4}{c}{\overline{S}^{\;<}} & &
\end{array}$$

Combinatory grammars also include type-raising rules, which turn arguments into functions over functions-over-such-arguments. These rules allow arguments to compose, and thereby take part in coordinations like *I dislike, and Mary likes, musicals*. For example, the following rule:

(52) *Subject Type-raising:* $(>\mathbf{T})$
$$NP : a \quad \Rightarrow \quad \mathrm{T}/(\mathrm{T}\backslash NP) : \mathbf{T}a$$

—or equivalently:

(53) *Subject Type-raising:* $(>\mathbf{T})$
$$NP : a \quad \Rightarrow \quad \mathrm{T}/(\mathrm{T}\backslash NP) : \lambda f.f\ a$$

—allows the conjuncts to form as in (54) below (again, the remainder of the derivation is given in the briefer notation):

(54)
$$\begin{array}{cccccc}
\text{I} & \text{dislike} & \text{and} & \text{Mary} & \text{likes} & \text{musicals}\\
\overline{NP} & \overline{(S\backslash NP)/NP} & \overline{CONJ} & \overline{NP : mary'} & \overline{(S\backslash NP)/NP : \lambda x.\lambda y.like'xy} & \overline{NP}\\
\overline{S/(S\backslash NP)}^{\;>\mathbf{T}} & & & \overline{S/(S\backslash NP) : \lambda f.f\ mary'}^{\;>\mathbf{T}} & & \\
\multicolumn{2}{c}{\overline{S/NP}^{\;>\mathbf{B}}} & & \multicolumn{2}{c}{\overline{S/NP : \lambda x.like'x\ mary'}^{\;>\mathbf{B}}} & \\
\multicolumn{5}{c}{\overline{S/NP}^{\;<\Phi>}} & \\
 & & \multicolumn{4}{c}{\overline{S}^{\;>}}
\end{array}$$

Rule (52) has an "order-preserving" property. That is, it turns the NP into a *rightward* looking function over *leftward* function, and therefore preserves the linear order of subjects and predicates.

Like composition, type-raising rules are required by the Principle of Combinatory Transparency (48) to be transparent to semantics. This fact ensures that the raised subject NP has an appropriate interpretation, and can compose with the verb to produce a function that can either coordinate with a transitive verb or reduce with an object *musicals* to yield $like'\ musicals'\ mary'$.

Since complement-taking verbs like *think*, VP/S, can in turn compose with fragments like *Mary likes*, S/NP, we correctly predict the fact that right-node raising is unbounded, as in a, below, and also provide the basis for an analysis of the similarly unbounded character of leftward

[19]The analysis begs some syntactic and semantic questions about the coordination rule and the interpretation of modals. See Steedman (1990, 2000a) for more complete accounts of both.

extraction, as in b (see the earlier papers and Steedman 1991, 1997 for details, including ECP effects and other extraction asymmetries, and the involvement of similar fragments in intonational phrasing):

(55) a. [I dislike]$_{S/NP}$ and [you think Mary likes]$_{S/NP}$ musicals.
  b. The musicals which [you think Mary likes]$_{S/NP}$.

This apparatus has been applied to a wide variety of coordination phenomena, including English "argument-cluster coordination", "backward gapping" and verb-raising constructions in Germanic languages, and English gapping. The first of these is relevant to the present discussion, and is illustrated by the following analysis, from Dowty (1988), itself related to the analysis of Dutch argument cluster coordination in Steedman 1985:[20]

(56)

$$\begin{array}{lllll}
\text{introduce} & \text{Bill} & \text{to Sue} & \text{and} & \text{Harry to George} \\
\hline
(VP/PP)/NP & (VP/PP)\backslash((VP/PP)/NP)\ {<}\mathbf{T} & VP\backslash(VP/PP)\ {<}\mathbf{T} & CONJ & VP\backslash((VP/PP)/NP)\ {<}\mathbf{B} \\
 & \multicolumn{2}{c}{VP\backslash((VP/PP)/NP)\ {<}\mathbf{B}} & & \\
 & \multicolumn{4}{c}{VP\backslash((VP/PP)/NP)\ {<}\Phi{>}} \\
\multicolumn{5}{c}{VP\ {<}}
\end{array}$$

The important feature of this analysis is that it uses "backward" rules of type-raising $<\mathbf{T}$ and composition $<\mathbf{B}$ that are the exact mirror-image of the two "forward" versions introduced as examples (47) and (52), which similarly guarantee that the semantics of non standard constituents like *Bill to Sue* is such as to reduce appropriately with a ditransitive verb like *introduce*. It is in fact a prediction of the theory that such a construction can exist in English, and its inclusion in the grammar requires no additional mechanism whatsoever.

The earlier papers show that no *other* non-constituent coordinations of dative-accusative NP sequences are allowed in any language with the English verb categories, given the assumptions of CCG. Thus the following are ruled out in principle, rather than by stipulation:

(57) a. *Bill to Sue and introduce Harry to George
  b. *Introduce to Sue Bill and to George Harry

In English the phenomenon shows up in all constructions that can be assumed to involve multiple arguments of the same functor:[21]

(58) a. I gave Deadeye Dick a sugar stick, and Mexican Pete a bun.
  b. I saw Keats yesterday, and Chapman the day before.
  c. I saw Gilbert arrive and George leave.
  d. I persuaded Warren to take a bath and Dexter to have a wash.
  e. I promised Mutt to go to the movies and Jeff to go to the play.
  f. I told Shem I lived in Edinburgh and Shaun I lived in Philadelphia.
  g. I bet Sammy sixpence I would win and Rosie a dollar I would lose.

[20]In more recent work, Dowty has disowned this analysis, on grounds that it entails an "intrinsic" use of logical form to account for binding phenomena. This issue is discussed in Steedman 1996, where it is argued that this use of logical form is no more "intrinsic" than that implicit in the separation of linear precedence from dominance in the "liberation categorial grammars" of Zwicky (1986), Hoeksema (1991), and Dowty (1996), and the related approach ofMuskens (1999).

[21]This assumption precludes a small clause analysis of the basic constructions.

A number of related well-known cross-linguistic generalizations first noted by Ross (1970) concerning the dependency of so-called "gapping" upon lexical word-order are also captured (as was pointed out in Dowty (1988), written in 1985, and by Steedman (1985, 1990, 2000a)). The pattern is that in languages whose basic clause constituent order subject-verb-object (SVO), the verb or verb group that goes missing is the one in the right conjunct, and not the one in the left conjunct. The same asymmetry holds for VSO languages like Irish. However, SOV languages like Japanese show the opposite asymmetry: the missing verb is in the *left* conjunct.[22] The pattern can be summarized as follows for the three dominant constituent orders (asterisks indicate the excluded cases):[23]

(59) SVO: *SO and SVO   SVO and SO
VSO: *SO and VSO   VSO and SO
SOV: SO and SOV *SOV and SO

This observation can be generalized to individual constructions within a language: just about any construction in which an element apparently goes missing preserves canonical word order in an analogous fashion: (56) above is an example of this generalization holding of a verb-initial construction in English.

Phenomena like the above immediately suggest that all complements of verbs bear type-raised categories in all languages. However, we do not want anything *else* to type-raise. In particular, we do not want raised categories to raise again, or we risk infinite regress in our rules. One way to deal with this problem is to explicitly restrict the two type-raising rules to the relevant arguments of verbs, as follows, a restriction that is a natural expression of the resemblance of type-raising to some generalized form of (nominative, accusative, etc) grammatical *case*—see Steedman 1985, 1990.

(60) *Forward Type-raising:* ($>\mathbf{T}$)
$X : a \quad \Rightarrow \quad \mathrm{T}/(\mathrm{T}\backslash X) : \lambda f.fa$
where $X \in \{NP\}$

(61) *Backward Type-raising:* ($<\mathbf{T}$)
$X : a \quad \Rightarrow \quad \mathrm{T}\backslash(\mathrm{T}/X) : \lambda f.fa$
where $X \in \{NP, PP, AP, VP, VP', S, S'\}$

The other solution is to simply expand the lexicon by incorporating of the raised categories that these rules define, so that categories like NP have raised categories, and all functions into such categories, like determiners, have the category of functions into raised categories.

These two tactics are essentially equivalent, because in some cases we need both raised and unraised categories for complements. (The argument is developed in Steedman 1997, and depends upon the observation that any category that is not a barrier to extraction must bear an

[22]A number of apparent exceptions to Ross's generalization have been noted in the literature, including Zapotec (which is VSO but allows SO and VSO; see Rosenbaum 1977) and German (which is SOV but allows SOV and SO; see van Oirsouw 1987). These are discussed in Steedman 2000a which shows how they are made possible by the fact that Zapotec (unlike Irish) allows SOV as a main-clause word order, while German and Dutch (unlike Japanese) allow VSO/SVO as main-clause orders. Ross's constraint is there stated in terms of overall order properties of languages and constructions rather than any notion of "underlying" word order.

[23]Languages that order object before subject are sufficiently rare as to apparently preclude a comparable data set, although any result of this kind would be of immense interest.

unraised category, and any argument that can take part in argument-cluster coordination must be raised). The correct solution from a linguistic point of view, inasfar as it captures the fact that some languages appear to lack certain unraised categories (notably *PP* and S'), is probably the lexical solution. However the restricted rule-based solution makes derivations easier to read and causes them to take up less space.

5 The Logic and Neurology of Language and Action

The ubiquitous appearance of composition **B** and type-raising **T** in both affordance-mediated action planning of the most elementary sort on the one hand, and universal grammar on the other, strongly suggests that the language faculty in its syntactic aspect is directly hung onto a more primitive set of prelinguistic operations including these combinators, originally developed for motor planning. This hypothesis has strong implications for the theory of evolution and the child's acquisition of language, for which there is considerable circumstantial evidence from neurological and neuroanatomical observations.

In particular, it has long been known that Broca's area, or the left inferior frontal area which evidence from brain imaging and acquired aphasias suggests is implicated in morphosyntactic processing, is immediately adjacent to areas involved in motor planning, suggesting that in evolutionary and developmental terms, the former are built upon the latter (see Greenfield et al. (1972); Greenfield (1991)).

For example, the onset in the child of the ability to compose motor plans such as those needed for composite reaching around an obstacle precedes the onset of productive language use. It is also argued by Deacon (1988, 1997) and Diamond (1990) to depend in evolutionary and developmental terms upon the mastery of response inhibition mediated by more frontal areas that are also implicated in language disorders. (It is interesting to note in this connection that lasting Broca's aphasia is generally associated with damage to such frontal areas, as well as to Broca's area itself.) Further suggestive support for the idea that such processing is closely related to action representations of the kind discussed in earlier sections comes from the fact that Broca's aphasics often show a specific deficit in accessing *verbs* in comparison to nouns (see Miceli et al. 1984, 1988 and Caramazza and Hillis 1991).

It has also been known since the pioneering work of de Laguna (1927) and Bruner (1968) that the onset of language in infants—and the entire cognitive explosion into the Piagetian operational phase—follows closely on the mastery of motor planning involving the use of tools at the final sixth stage of the Piagetian sensory-motor phase of cognitive development.

The Linear Dynamic Event Calculus and related linear and STRIPS-like systems offer a way of representing actions in ways that are useful for planning action. This in turn offers a way of capturing affordances of objects, a notion that is again relevant to efficient planning in the real world. Two combinatory operations of composition and type-raising play a central role in this process. Those same combinators appear in syntactic guise in natural language, where they provide the basis for an explanatory account of language-specific constructions and cross linguistic universal generalization, and where there is circumstantial evidence from neuroanatomy and child development that motor planning and language are closely related. It has always seemed likely that syntax is related in evolutionary terms to a prelinguistic faculty related to planned action (Lashley 1951; Miller et al. 1960). LDEC and CCG make that relation look direct enough to explain the fact that the evolutionary advance in question appears to have been very rapid indeed.

It is interesting to speculate upon what such an evolutionary step might be based. One candidate that has considerable theoretical appeal is the attainment of the modal and propositional attitude concepts that are required in order to reason about other minds—that is, functions over propositional entities. (We have so far glossed over an important distinction between plans, which compose actions of type *state* → *state*, and grammar, which composes functions of type *proposition* → *proposition* or *property* → *property*.)

It is propositional functions that induce true recursion in both conceptual structures and grammar. There is no evidence that apes entertain concepts of the requisite kind. In particular, the most successful attempts to teach apes to use language, notably those involving ASL and other manipulative languages (Gardner and Gardner 1969, Premack 1971, 1986), show a lack of recursive syntax coupled with an almost autistic paucity of goal-directed conversational initiative. Perhaps a theory of other minds and the associated propositional attitude concepts are *all* that is missing, consistent with proposals by Tomasello (1999).

The interest of these observations lies in the likely relation of the infants first propositional concepts to speech acts. Speech acts are actions like any other and can be represented in the kind of event calculus considered here. However, the entities that they add to the database are propositional. It is likely that the key to this phase of infant language development and its divergence from that of other primates is to be found in the transition between making simple motor plans of the kind investigated by Diamond (1990) up to month 12, and making plans involving other agents' actions, of the kind investigated by de Laguna (1927), Bruner (1968), and authors collected in Bullowa (1979).

Acknowledgments

Earlier versions of this paper were presented at the Gatsby Computational Neuroscience Unit, University College London, Nov. 2000 and the Maryland Mayfest 2001: Cognitive Neuroscience of Language. May, 2001. Thanks to the participants and to Ric Alterman, Joyce McDonough, Silvia Gennari, Peter O'Hearn, Mark Johnson, Murray Shanahan, Matthew Stone, and Rich Thomason for comments and advice on drafts. Various stages of the research were funded in part by EPSRC grants GR/M96889 and GR/R02450 and EU (FET) grant MAGICSTER.

References

Ades, A. and Steedman, M. (1982). On the order of words. *Linguistics and Philosophy*, 4:517–558.

Armelín, P. and Pym, D. (2001). Bunched logic programming. In *Proceedings of IJCAR 2001*. to appear.

Bach, E. (1983). Generalized categorial grammars and the english auxiliary. In Heny, F. and Richards, B., editors, *Linguistic Categories: Auxiliaries and Related Puzzles, II*, pages 101–120. Reidel, Dordrecht.

Barker, C. (2001). Continuations and the nature of quantification. *Linguistics and Philosophy*. to appear.

Bibel, W., del Cerro, L. F., Fronhfer, B., and Herzig, A. (1989). Plan generation by linear proofs: on semantics. In *German Workshop on Artificial Intelligence - GWAI'89*, volume 216 of *Informatik-Fachberichte*, Berlin. Springer Verlag.

Bruner, J. (1968). *Processes of Cognitive Growth: Infancy*. Clark University Press, Worcester, MA.

Bullowa, M., editor (1979). *Before Speech: the Beginning of Interpersonal Communication*. Cambridge University Press, Cambridge.

Caramazza, A. and Hillis, A. (1991). Lexical organization of nouns and verbs in the brain. *Nature*, 349:788–790.

Cervesato, I., Hodas, J. S., and Pfenning, F. (1996). Efficient resource management for linear logic proof search. In Dyckhoff, R., Herre, H., and Schroeder-Heister, P., editors, *Proceedings of the 1996 Workshop on Extensions to Logic Programming*, number 1050 in Lecture Notes in Artificial Intelligence, pages 67–82. Springer-Verlag.

Curry, H. B. and Feys, R. (1958). *Combinatory Logic: Vol. I.* North Holland, Amsterdam.

de Laguna, G. (1927). *Speech: Its Function and Development.* Yale University Press, New Haven, CT.

Deacon, T. (1988). Human brain evolution i: Evolution of human language circuits. In Jerison, H. and Jerison, I., editors, *Intelligence and Evolutionary Biology*. Springer-Verlag, Berlin.

Deacon, T. (1997). *The Symbolic Species*. Norton, New York.

Diamond, A. (1990). Developmental time course in human infant and baby monkeys and the neural bases of inhibitory control in reaching. In Diamond, A., editor, *The Development and Neural Bases of Higher Cognitive Functions*, pages 637–676. New York Academy of Sciences, New York.

Dowty, D. (1972). *Studies in the Logic of Verb Aspect and Time Reference in English*. PhD thesis, University of Texas at Austin.

Dowty, D. (1979). Dative movement and thomason's extensions of montague grammar. In Davis, S. and Mithun, M., editors, *Linguistics, Philosophy, and Montague Grammar*, pages 153–222. University of Texas Press, Austin, TX.

Dowty, D. (1988). Type-raising, functional composition, and nonconstituent coordination. In Oehrle, R. T., Bach, E., and Wheeler, D., editors, *Categorial Grammars and Natural Language Structures*, pages 153–198. Reidel, Dordrecht.

Dowty, D. (1996). Towards a minimalist theory of syntactic structure. In Bunt, H. and van Horck, A., editors, *Discontinuous Constituency*, pages 11–62. Mouton de Gruyter, The Hague.

Fikes, R. and Nilsson, N. (1971). Strips: a new approach to the application of theorem proving to problem solving. *AI Journal*, 2:189–208.

Gardner, R. A. and Gardner, B. (1969). Teaching sing language to a chimpanzee. *Science*, 165:664–672.

Gelfond, M. and Lifschitz, V. (1993). Representing actions and change by logic programs. *Journal of Logic Programming*, 17:301–321.

Gibson, E. and Spelke, E. (1983). The development of perception. In Mussen, P., editor, *Handbook of Child Psychology, vol. 3: Cognitive Development*, pages 1–76. Wiley, New York.

Gibson, J. (1950). *The Perception of the Visual World*. Houghton-Mifflin Co., Boston, MA.

Gibson, J. (1966). *The Senses Considered as Perceptual Systems.* Houghton-Mifflin Co., Boston, MA.

Gibson, J. (1979). *The Ecological Approach to Visual Perception*. Houghton-Mifflin Co., Boston, MA.

Girard, J.-Y. (1987). Linear logic. *Theoretical Computer Science*, 50:1–102.

Girard, J.-Y. (1995). Linear logic: its syntax and semantics. In Grirard, J.-Y., Lafont, Y., and Regnier, L., editors, *Advances in Linear Logic*, volume 222 of *London Mathematical Society Lecture Notes*, pages 1–42. Cambridge University Press, Cambridge.

Greenfield, P. (1991). Language, tools and brain. *Behavioral and Brain Sciences*, 14:531–595.

Greenfield, P., Nelson, K., and Saltzman, E. (1972). The development of rule-bound strategies for manipulating seriated cups: a parallel between action and grammar. *Cognitive Psychology*, 3:291–310.

Hanks, S. and McDermott, D. (1986). Default reasoning, nonmonotonic logics, and the frame problem. In *Proceedings of the 8th National Conference on Artificial Intelligence (AAAI-86)*, pages 328–333.

Harel, D. (1984). Dynamic logic. In Gabbay, D. and Guenthner, F., editors, *Handbook of Philosophical Logic*, volume II, pages 497–604. Reidel, Dordrecht.

Hoeksema, J. (1991). Complex predicates and liberation in dutch and english. *Linguistics and Philosophy*, 14:661–710.

Hölldobler, S. and Sneeberger, J. (1990). A new deductive approach to planning. *New Generation Computing*, 8:225–244.

Japaridze, G. (1998). *The Logic of Resources and Tasks*. PhD thesis, University of Pennsylvania.

Köhler, W. (1925). *The Mentality of Apes*. Harcourt Brace and World, New York.

Kowalski, R. (1979). *Logic for Problem Solving*. North Holland, Amsterdam.

Kowalski, R. and Sergot, M. (1986). A logic-based calculus of events. *New Generation Computing*, 4:67–95.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things*. University of Chicago Press, Chicago.

Lambek, J. (1958). The mathematics of sentence structure. *American Mathematical Monthly*, 65:154–170.

Lashley, K. (1951). The problem of serial order in behavior. In Jeffress, L., editor, *Cerebral Mechanisms in Behavior*, pages 112–136. Wiley, New York. reprinted in Saporta (1961).

Lee, D. (1980). The optic flow field: The foundation of vision. *Philosophical Transactions of the Royal Society, Series B*, 290:169–179.

Martí-Oliet, N. and Meseguer, J. (1999). Action and change in rewriting logic. In Pareschi, R. and Fronhöfer, B., editors, *Dynamic Worlds*, pages 1–53. Kluwer, Dordrecht.

McCarthy, J. and Hayes, P. (1969). Some philosophical problems from the standpoint of artificial intelligence. In Meltzer, B. and Michie, D., editors, *Machine Intelligence*, volume 4, pages 473–502. Edinburgh University Press, WEdinburgh.

Miceli, G., Silveri, M., Nocentini, U., and Caramazza, A. (1988). Patterns of dissociation in comprehension and production of nouns and verbs. *Aphasiology*, 1:351–358.

Miceli, G., Silveri, M., Villa, G., and Caramazza, A. (1984). On the basis for the agrammatic's difficulty in producing main verbs. *Cortex*, 20:207–220.

Miller, G., Galanter, E., and Pribram, K. (1960). *Plans and the Structure of Behavior*. Holt, London.

Moens, M. (1987). *Tense, Aspect, and Time Reference*. PhD thesis, Edinburgh.

Moens, M. and Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics*, 14:15–28.

Moortgat, M. (1997). Categorial type logics. In van Benthem, J. and ter Meulen, A., editors, *Handbook of Logic and Language*, pages 93–177. North Holland, Amsterdam.

Muskens, R. (1999). Underspecified semantics. In von Heusinger, K. and Egli, U., editors, *Reference and Anaphoric Relations*, pages 311–338. Kluwer, Dordrecht.

Narayanan, S. (1997). Talking the talk *is* like walking the walk: A computational model of verbal aspect. In *Proceedings of the Nineteenth Conference of the Cognitive Science Society, Stanford, August*.

Narayanan, S. (1999). Reasoning about actions in narrative undertanding. In *Proceedings of the International Joint Conference on Artificial Intelligence, Stockolm, August*, pages 350–358, San Francisco, CA,. Morgan Kaufmann.

Norman, D. (1988). *The Psychology of Everyday Things*. Basic Books, New York.

Norman, D. (1999). Affordance, conventions, and design. *Interactions*, 6:38–43.

O'Hearn, P. and Pym, D. (1999). The logic of bunched implications. *Bulletin of Symbolic Logic*, 5:215 – 244.

Premack, D. (1971). Language in the chimpanzee? *Science*, 172:808–822.

Premack, D. (1986). *Gavagai!* MIT Press, Bradford Books, Cambridge, MA.

Pym, D. (1999). On bunched predicate logic. In *Proceedings of the 14th Symposium on Logic In Computer Science (LICS)*. IEEE Computer Society.

Pym, D. (2001). *The Semantics and Proof Theory of the Logic of Bunched Implications*. to appear.

Rescorla, R. and Wagner, A. (1972). A theory of pavlovian conditioning. In Black, A. and Prokasy, W., editors, *Classical Conditioning, II*. Appleton-Century-Crofts, New York.

Rosenbaum, H. (1977). Zapotec gapping as counterevidence to some universal proposals. *Linguistic Inquiry*, 8:379–395.

Ross, J. R. (1970). Gapping and the order of constituents. In Bierwisch, M. and Heidolph, K., editors, *Progress in Linguistics*, pages 249–259. Mouton, The Hague.

Sandewall, E. (1994). *Features and Fluents*. Clarendon Press, Oxford.

Saporta, S., editor (1961). *Psycholinguistics: A Book of Readings*. Holt Rinehart Winston, New York.

Schoppers, M. J. (1987). Universal plans for reactive robots in unpredictable environments. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence (IJCAI 87)*, pages 1039–1046, Milan, Italy.

Schubert, L. (1994). Explanation closure, action closure, and the sandewall test suite for reasoning about change. *Journal of Logic and Computation*, 4:679–700.

Shanahan, M. (1997). *Solving the Frame Problem.* MIT Press, Cambridge.

Shieber, S. (1986). *An Introduction to Unification-Based Approaches to Grammar*. CSLI Publications, Stanford, CA.

Smullyan, R. (1985). *To Mock a Mockingbird*. Knopf, New York.

Smullyan, R. (1994). *Diagonalization and Self-Reference.* Clarendon Press, Oxford.

Steedman, M. (1985). Dependency and coordination in the grammar of dutch and english. *Language*, 61:523–568.

Steedman, M. (1987). Combinatory grammars and parasitic gaps. *Natural Language and Linguistic Theory*, 5:403–439.

Steedman, M. (1990). Gapping as constituent coordination. *Linguistics and Philosophy*, 13:207–263.

Steedman, M. (1991). Structure and intonation. *Language*, 67:262–296.

Steedman, M. (1995). Dynamic semantics for tense and aspect. In *Proceedings of the International Joint Conference on Artificial Intelligence, Montreal, Quebec*, pages 1292–1298, San Francisco, CA. Morgan Kaufmann.

Steedman, M. (1996). *Surface Structure and Interpretation*. MIT Press, Cambridge, MA.

Steedman, M. (1997). Temporality. In van Benthem, J. and ter Meulen, A., editors, *Handbook of Logic and Language*, pages 895–938. North Holland, Amsterdam.

Steedman, M. (2000a). *The Syntactic Process*. MIT Press, Cambridge, MA.

Steedman, M. (2000b). *The Productions of Time*. Ms., University of Edinburgh, `http://www.cogsci.ed.ac.uk/~steedman/`.

Steedman, M. (2002). Formalizing affordance. In *Proceedings of the 24th Annual Meeting of the Cognitive Science Society, Washington D.C.*, Mahwah NJ. Lawrence Erlbaum. to appear.

Thielscher, M. (1998). Reasoning about actions: Steady versus stabilizing state constraints. *Artificial Intelligence*, 104:339–355.

Thielscher, M. (1999). From situation calculus to fluent calculus: State update axioms as a solution to the inferential frame problem. *Artificial Intelligence*, 111:277–299.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press, Cambridge, MA.

Turvey, M. (1990). Coordination. *American Psychologist*, 45:938–953.

van Benthem, J. (1991). *Language in Action*. North Holland, Amsterdam.

van Oirsouw, R. (1987). *The Syntax of Coordination*. Croom Helm, London.

Vendler, Z. (1967). *Linguistics in Philosophy*. Cornell University Press, Ithaca, NY.

Verkuyl, H. (1972). *On the Compositional Nature of the Aspects*, volume 15 of *Foundations of Language Supplementary Series*. Reidel, Dordrecht.

Wood, M. M. (1993). *Categorial Grammar*. Routledge, London.

Young, R. and Morgan, W. S. (1987). *The Navaho Language*. University of New Mexico Press, Albuquerque, NM.

Zwicky, A. (1986). Concatenation and liberation. In *CLS 22: Papers from the General Session*, Chicago. Chicago Linguistic Society, University of Chicago.