

Deconstructing Functionalist Explanations of Linguistic Universals

Still glides the Stream, and shall forever glide;
The Form remains, the Function never dies.
Wordsworth, 1820

1.0 Introduction

The more we learn about what syntax is, the less we know about why it is the way it is. “Functionalist” interpretations offer potential explanations for the peculiarities of syntactic structure and characteristically blend what they are with why they exist. In this paper, I review a few examples of such explanations, with several morals in mind. First, it is critical that functional explanations of linguistic structures have independent theoretical and empirical support; second, functional explanations can expand beyond the traditional applications, which are most often based on the communicative needs and semantic habits of speakers. Third, we may be quite surprised to realize that what starts out as a cognitive explanation of a linguistic property may reverse, and show us how a linguistic property might explain the cognitive property.

Functionalism has become a controversial topic in linguistic theory (see e.g., Newmeier, 1998, for an excellent review). Around it swirl disputes over whether grammar is a formally independent mental structure, and if so, whether it is computationally arbitrary. The functionalist answers to such questions are characteristically, (1) that grammatical structures either are an ephemeron, or (2) that they are caused and shaped historically by the functional role of language. This controversy has been attached to larger issues in cognitive science, such as the notion that the mind is organized in modules – if grammar is independent of other systems, it is a quintessential example of a module. If it is interdependent and not self-defined, then one of the strongest prima facie cases of modularity falls aside.

I start with some general consideration of functionalist explanations, drawn from examination of a very simple communication system. Then, I review three cases, which exemplify functional constraints on linguistic structures of different kinds; one from language perception, one from the biological underpinnings of language ability, and one from language learning. All three involve an aspect of how sentences are composed of ordered sequences that express hierarchical structures. In each case, I will question whether the particular linguistic property is the result of extra-linguistic constraints, or conversely, whether the computational requirement of mapping serial form and hierarchical structure unique to language itself underlies more general constraints on human cognition. My ambitious goal is to convince even the most dogmatic nativist formalist linguo-centrist linguist that eclectic consideration of functionalist interpretations of linguistic universals can be illuminating in surprising and helpful ways.

2.0 Meaning and language go to the dogs

Language has a number of apparent functions. These include communication of meanings to others and representation of meanings to oneself. These roles set constraints on the nature of language and the useful structure of meanings.

Which is logically prior, meaning or language? We ask this because we want to have some idea which kind of system is setting the groundwork constraints. This question is not a complete “chicken-and-egg” conundrum; we do have some evidence about communication which suggests that potentially shared meanings are independent of language.

An example of this is how what I can communicate with my poodle, Cassie, determines our mutual language (Limber, 1995, personal communication). Cassie and I don't have much to say to each other about most of the world, we simply view it together: her primary fascination is other dogs, lesser animals, and the smells they have left behind, all of little interest to me. She clearly has theories about the world that enter some domains I can only dimly sense (e.g., visual aspect of other dogs), and others I cannot sense at all (myriad smells in the air, grass, trees). Conversely, her perceptual system limits her ability to discriminate human artifacts – even though it would help her greatly, for example, in getting into the right car in a parking lot. But we do have certain fundamental concepts that we share: water, food, really good food, walk, car, stay, sit, down, (find)addie (a puppyhood co-puppy), bone, puppy, (find)LouAnn (a former co-owner), (get)down, (get)up, bark, (go)out, (go)to my-office, (go)in, (get)off, (go in) house, (get)on, (get)the ball, no, ok. This may seem like a lot for a non-circus dog, but one must remember that she is a poodle.

The questions at hand are: is our communication system a language? How is it constrained by our mutual meaning systems? I will take the unconventional position (for linguists), that Cassie and I do share a language, albeit one that is so heavily constrained by minimal cognitive overlap, that it remains very simple. The fundamental fact is that we do communicate effectively about those things that we actually share semantically. If we shared more semantically, I see no principled upper limit on the number of different “words” we could use. If, for example, I could differentiate smells the way Cassie can, this would enrich our discussions enormously, at least from her standpoint. The same would be true if she could differentiate cars – our life in parking lots would be a great deal more communicative and easier, at least from my standpoint. In the long theoretical run, it will matter a great deal what notion of “semantics” Cassie and I share. One might argue that it is not “intensions” or propositional attitudes, but only extensions, aka referents. The fact that we share variable but appropriate behaviors in relation to each concept suggests some kind of shared propositional representation. But I will not argue that here, because it certainly requires a lot of demonstration. The essential point is that whatever the true nature of our shared concepts, they are the only things that we can convincingly communicate about.

Of course, it may be that our communication system is ALSO constrained by Cassie's (and my) computational limitations. In general, she does not appear to respond to combinations of words, and so does not exhibit much of what we usually call "syntax", the system that governs how word elaborations and combinations are formed. An exception is the negative operator, "no", which can be combined with actions or expectations she has. She can be told to stop barking, or that she is NOT going in the car, or that she is NOT going to get a treat. A critical (informal) experiment on the productivity of "no" involves a neutral action, such as "sit": she will refrain from sitting, if "sit" and "no" are presented together. (Though she does not look happy about the situation). Thus, "no" does not merely mean to her that something bad will happen.

But, it is not clear that these are syntactic combinations bound into larger units of meaning, as opposed to a serial presentation of information in which each element modifies what it is adjacent to. In practical terms, does word order affect Cassie's interpretation of "no"? In fact, it does: in Cassie's "syntax", "no" has to follow what it negates, not precede it: thus, if I say, "no, bark", she responds to "no" by looking guilty, but continuing to bark. If I say, "bark, no", then she stops barking. If I say "no, car" she clearly expects (after briefly looking crestfallen) a ride in the car: but, if I say "car, no", she gets the disappointing picture right away. Thus, she evidences understanding of the frame, "X, no" with enough statistical regularity to satisfy those experimental psychologists who allegedly teach dolphins, sea-lions and chimps to talk. So, we might surmise that Cassie has a syntactic rule: "no" must follow what it modifies.

Now the question is: can we explain Cassie's syntactic constraint that "no" follows its argument by reference to another aspect of Cassie's cognition. One can surmise that "no" needs an already established mental representation to apply to, so it works out best if I elicit a representation in her before she hears the negation that applies to it. (An alternative interpretation is that dogs use Polish notation (in which operators follow their arguments), or at least poodles do). In this sense, she is not combining "no" in a syntactic frame at all; rather, she is responding serially to representations as she hears them. On this view, Cassie cannot hold operators such as negation in mind without something to operate on, but objects or action-ideas can be held in mind on their own. This is not an unintuitive idea, but, to go through in a profound way, we need other evidence that dogs (or poodles, anyway) think serially and allow the contents of a current representation to modify a prior one. However, let us assume that such evidence is forthcoming (certainly, most dog owners would attest to it). Then we would have an example of a functional explanation of what might otherwise be taken to be a syntactic process. Cassie has to hear "no" after the idea it applies to, because she could not hold it in mind if it preceded the idea it applies to.

Further evidence for this kind of explanation comes from Cassie's treatment of the general "find" instruction, "where is X?" In this case, the operator must precede

the argument. If I say, “puppy, where?”, Cassie merely looks puzzled; but if I say, “where puppy?”, she goes to the window, looks out and barks. This pattern, too, generalizes across goals. If we think of this in terms of the underlying semantic operations, we can make sense of why it is different from the negative operator. The “where is” or “find” instruction can be acted on before hearing the argument: indeed, if I say just “Where?....” Cassie looks alert and may start running around, waiting for the orienting information. Thus, “where is” has an immediate enactable referent. Another way of putting this is that “where” is a modal modification of the entire utterance, as opposed to an operator directly on the argument. Here too, rather than postulating a syntactic frame, we can interpret Cassie’s understanding in terms of how the utterance functions.

The reader may think all of this is a *reductio ad canem*, on the topic of functional explanations of ‘syntax’, and there may be something to that. I am personally convinced of what I claim about Cassie, but I would agree that it requires more scientific justification and breadth to be taken as serious Dog Science.

We can, however, derive some principled morals on how to think about functionalist explanations for syntactic processes. If one is uncomfortable with my reasoning above, it could be because I did not develop an *independent* demonstration of the idea that Cassie cannot hold a semantic operator in mind, but must apply it to an already established representation. That of course, can be tested experimentally, for example with directional or negation operators presented before the objects to which they apply in a standard animal learning paradigm. The same argument and response goes for the locative imperative. Until the principles are verified in general, it rests on intuition alone, and the appearance of being *post hoc*. I also did not demonstrate the natural priority of shared semantic concepts in what we communicate about, I merely asserted it. And, of course, I did not prove that Cassie’s sensitivity to word order is compositionally syntactic in any sense except superficially.

So, to be fully satisfactory, a functionalist explanation is most convincing when it has the following characteristics:

- a) The property to be explained is truly syntactic.
- b) There is independent evidence for the explanatory functional principle beyond its effect on language.

Most important, the result of this consideration is to make clear that a functional explanation may not be adequate when merely referring to “meaning” as determining form. In this case, the putative functional explanation for apparent constraints on form, have to do with how the meaning is processed, not just what it is.

But there is an even more profound question. I started with the assumption that Cassie and I share certain meanings in our mutual cognition, and that this limits

and focuses the kind of language we share. However, how can we be sure that the causal link is not in the reverse direction? That is, is our apparent shared cognition itself shaped by our shared symbolic system? The ultimate answer to this will depend on far more serious study of canine cognition, and more success in understanding human cognition than we currently have. At the moment, we do not have a strong theoretical basis to justify the intuition that dogs and humans start with shared cognition and build language around it. And, indeed, it may turn out to be the other way.

3. Case studies in human/human language

Let's see how these principles guide us in considering a range of alleged functional explanations of syntactic universals.

3.1 Me first! The order of thought, and the thought of order.

It is often noted that languages tend to present agents prior to patients. Some languages appear to require this grammatically under certain circumstances. Recently, Aissen and Bresnan (in press) and Bresnan (2001) have attempted to integrate a constraint within a stochastic optimality-theoretic framework. On this view, Agent->Patient is a linguistic principle, embedded as a constraint, which will appear as grammatical in certain languages, if it is ranked high enough in the overall ranking of constraints. If it is not ranked very high, it can still have occasional effects, because of its probabilistic penumbra, which occasionally elevates it to a high rank. This treatment appears to bridge with one principle, the fact that speakers tend to prefer sentences with Agent->Patient order and the generally observed fact that if a language grammaticizes order, that will be the order required.

Is this truly a linguistic principle, i.e., a universally present formal constraint on possible languages, or can it be explained by reference to some aspect of how language functions and is learned?

In the middle of the last century, a noted pioneer cognitive psychologist, Jerome Bruner observed that languages seem to prefer sentence structures in which the "agent" precedes the "object". He suggested that this grows out of early language learning, especially language production: children have an idea, and they express it in serial parts that correspond to the temporal order of the idea's conception. Bruner reasoned that the child starts with sentences about hiser own plans, desires, etc. The child realizes that its own intentions precede its actions. Through this, the child discovers *intentions* more generally, and the knowledge that they definitionally precede their corresponding actions. Agents must be animate, verbs tend to be activated intentionally and objects tend to be inanimate: thus, agents will tend to appear before objects by virtue of the serial stages of the underlying thought processes.

This is an appealing idea, but it did not have independent empirical support. However, it has the virtue of making some predictions about when the Agent->Patient order might be less optimal or violated, and when it would be the strongest. In particular, when the Agent is animate and the patient is inanimate, the restriction should be weakest: intermediate when both are animate or both are inanimate, and strongest when the Agent is inanimate and the Patient is animate. Furthermore, since intentions are discovered first within the child, it is reasonable to expect that if there are cases in which animates are ordered in relation to each other, 1st person will have the highest rank: since utterances are generally to a second person, that will be the next highest rank, with third (often non-present) person the lowest rank.

Of course, we would be advised to have some independent empirical support for Bruner's thesis, intuitively appealing as it is. In fact, some of the underlying assumptions about the hypothesis do not have support. In particular, Bruner assumed that the Agent->Patient order was fundamental to the earliest stages of language, since it was based on a theory of how the infant thinks. This predicts that the order should appear in the earliest stages of language use, and build up from there. There is some evidence, going back to Lois Bloom (1970), that early utterances tend to maintain Agent->Verb, and Verb-Patient order. But of course, in English, that is the canonical order anyway, so it is ambiguous as to whether the child's basis for the order is the order of thought, or what s/he is hearing.

A different way to approach the development of the order is to examine children's comprehension strategies. In a series of studies, we did just that (Bever, Mehler, Valian and Epstein, reported in Bever, T. G. 1970). We tracked the child's comprehension of reversible actives, passives, subject clefts and object clefts (we developed a version of the "puppet acting out" task, in which children are the puppeteers). We found something very surprising: 2 year olds comprehend *all* of the sentence types better than chance, indeed they understand object-clefts as well as they understand actives! But 3-4 year olds understand passives and object clefts systematically *worse* than chance, while comprehension of actives and subject clefts approaches perfection. We concluded that this shows that the child *learns* the regularity that Agents precede Patients (in English), rather than creating it by virtue of the order of thought.

Similarly, we tested the development of the child's sensitivity to semantic constraints, using sentences like "the dog ate the cookie" vs. "the child was eaten by the cookie" etc. Even more surprising than the prior result, we found that two year old children were only slightly sensitive to the semantic constraints: they often correctly acted out anomalous sentences (e.g., by having the cookie bounce all over the dog, and announcing as one child did, "now the doggie inside cookie, that silly!"). Again it was older children who became selectively dependent on semantic reversibility. The generalization that Agents are animate, and Patients inanimate, is not available to the youngest children, it is *learned* as a strategy for sentence comprehension.

We later tested a variety of sentence orders and inflections on languages that have a range of each (notably Italian, Serbo-Croatian and Turkish, (Slobin, and Bever, 1982). We found that in each language, children develop comprehension strategies that are tuned to the contingencies in their language. Most notably, this meant that in Turkish, there was little attention to word order, but great attention to the presence and location of the object inflection marker. Most important, these strategies were learned, not the initial stage of comprehension.

The acquisition of such strategies is part of building up the processes of everyday sentence comprehension (and probably production). There is considerable research with adults showing that an early step in comprehension depends on such strategies, quite independent of the immediately following syntactic assignment of a derivation. (Townsend and Bever, 2001). Thus, in the end, the principles governing the usual relation between Agency, order and animacy are not part of the initial grammatical foundation of language. They are learned as part of learning to use language, as convenient (and probably necessary) shortcuts through syntactic computation. This does not change the validity of the claim that the principles have an extra-linguistic basis, it simply clarifies how the interaction between language behavior and the principles work. It does, however, give us empirical support to reject the claim that the principles are a direct part of the innate basis for language. Rather, on our view, they emerge out of the acquisition of the processes of language use, which automatically constrain the kinds of languages that can be learned.

The acquisition of strategies also plays a critical role in a dialectical model of language acquisition (Townsend and Bever, 2001). On this model, acquisition (logically) alternates between two kinds of acquisition processes: the abductive discovery of structural principles with limited domain, and the extension of that domain by the induction of extendable regularities in that domain. In the example given, this cashes out in English as:

- The development of simple declarative transitive sentence frames, based primarily on early acquisition of the verb+patient constituent, with agent status assigned by default to the other NP. Thus this is limited to short sentences with clearly locatable arguments.**
- The development of NVN the strategy based on the simple declarative order. This extends the domain of sentence production/understanding to longer utterances, with embeddings (themselves subject to the NVN constraint).**
- The development of exceptions to the NVN strategy as marked morphologically or in other ways. For example, object-first clefts and passives, which are systematically misunderstood during the development of the NVN strategy, become differentiated by virtue of their exhibited structural features.**

Consider the implications of this treatment for Aissen and Bresnan's ideas about the continuity of a discourse tendency and grammaticization for Agent->Patient. (Bresnan, and Aissen, (in Press), Bresnan 2001a, b). There are two aspects to their

ideas. On the one hand, they recognize a tendency across many languages (which certainly has been noted before), even when the order constraint is not a grammatical one. On the other hand, they offer stochastic optimality theory as a unifier of tendencies and grammatical requirements as a function of how important the constraint is in a given language. I leave to others the critique of how well and consistently optimality theory does the job of describing tendencies and grammatical restrictions (see, e.g., Nemeyer, 2002a,b). The issue here is the ontology of grammar that it represents. The virtue cited by Aissen and Bresnan is that the OT treatment is unified – one principle has statistical or grammatical effects depending on how highly it is ranked. Occam's razor is invoked.

This is fine if one has as a goal merely to describe the statistical and categorical facts with the least amount of ink. But if one aspires to propose that the grammar (or at least its architecture) is a model of human knowledge, then matters become bewildering. If I know that a given constraint has a specific frequency distribution that overlaps with other constraints, how does this cash out in my grammaticality intuitions? Why are our intuitions for specific sentences fairly constant, as opposed to stochastically wandering?

Aside from the explanatory virtues of the functionalist interpretation of the ordering priority of agents, we would like to have some "new" facts explained by some of the same general principles. That is, a functionalist interpretation of a phenomenon (like any theory) does best when it leads to other phenomena.

It appears that the general notion of an ordering hierarchy involving animacy, and centrality to a sentence has some currency in various ways across languages. The general idea is that more highly ranked phrases must precede less highly ranked phrases in all sentences, regardless of their functional role: special markers occur when the patient is the more highly ranked phrase. To quote Jelinek (personal communication)

"Navajo has an elaborate set of restriction on Agent/Patient order that are manifested in an Animacy Hierarchy and the famous Active/Inverse Voice alternation (the yi-/bi-) alternation. If the Patient is higher than the Agent on the animacy hierarchy, then it must come first, and the verb is in the Inverse voice. For example, you need to say the equivalent of

The boy was kicked by the horse.

using the Inverse, which is comparable to a Passive, except that it is a transitive, with two direct arguments (no "by" phrase). This kind of a hierarchy is also manifested in many Native American languages. For example, in Lummi (Salish) you can say

I hit him

but you can't say

he hit me

you have to say

I was hit (by him)

using the Passive. The rule is 1st, 2nd, 3rd person, NP. This kind of Agent hierarchy is very common.

This is part of a complex hierarchy for Navajo, which orders NPs: Adult human > child > animal > inanimate. Finally Willie (Willie, MaryAnn and Jelinek, Eloise. Manuscript), also noted that the hierarchy includes the constraint that topicalized definite nouns must precede indefinites. Jelinek points out that such hierarchies occur in many languages, even English has some such constraints for indirect objects and beneficiaries: only animate beneficiaries can occur prior to the benefit.

He mixed a drink for his friend
He mixed his friend a drink

He mixed some gasoline for his racecar.
*He mixed his racecar some gasoline

Put in more general terms, she writes:

Despite the surface diversity [in many different languages], argument hierarchies are alike in the following two traits:

1. There is some ranking of arguments according to a scale of PRESUPPOSITIONALITY. Some subdomain of the following ranking is observed (Jelinek 2000).

1,2>3>NPs> rankings within NPS according to definiteness, animacy, etc.

a. There is some SYNTACTIC manifestation of the hierarchy [which]....orders arguments in accord with the presuppositionality scale.

Finally, she sums up her overall generalization concerning ordering hierarchies:

Argument hierarchies are the morphosyntactic registration of....the Presuppositionality Scale....in particular languages....[they] represent the grammaticalization of certain highly frequent discourse tendencies.

If Jelinek is right, the alignment of an intuitive order hierarchy and Agent/Patient word order is part of the grammar of certain languages, while only preferred in others. Thus, one might argue that the fact that the constraint can be grammatical shows that it is truly part of linguistic universals, as Bresnan proposes. But, it is more complicated than just requiring agents to precede patients linearly. In certain languages, agents can follow patients, but then must be marked if that results in a linear order that violates the hierarchy. In both expressions, invoking the universal would not explain why the Agency/ordering hierarchy exists, nor why it has a characteristic order. Bruner's general hypothesis offers a true explanation of the

whole phenomenon. (Bruner, 1964). On this view, languages *could* have other kinds of constraints on Agency, and other orders of Agency hierarchies: but the language learning child naturally filters out those orders and constraints that do not conform to the order and processes of attributing the presupposed parts of sentences.

We can generalize this argument to the fuller order hierarchy which Jelinek summarizes. The sometimes-grammatical constraint that persons 1 > 2 > 3, animate > inanimate, definite > indefinite, are interpretable as a linguistic codification of conceptual hierarchies, which can differ in richness between cultures, but which have common themes. The surface order constraints differ from the Agent > Patient constraint, in the sense that it is the conceptual content of the phrases that governs the order rather than their functional syntactic role. However, the deeper consistency is that the priority of thought has a reflex in the precedence of the corresponding linguistic form. Jelinek's way of characterizing the basis for the hierarchy as "presuppositional" is particularly felicitous for our explanation: basically it proposes that the more presupposed an element is, by virtue of being a default in actions, the earlier it must appear in a sentence.

The hypothesis that the mapping from presupposition to linguistic constraint via discourse and perceptual strategies, leads to a rich research program. The idea is that the presuppositional hierarchy based on categories and/or discourse markers, reflects salience in conceptual access during speaking and listening. Salience can be reflected in surface order, which is exactly what happens. The mechanism for learning this in childhood is essentially automatic for conceptual categories and basic for syntactically marked presuppositionality (topic, focus, definiteness). For conceptual category based hierarchies it simply falls out of the conceptual system, apparently universal, possibly innate in an interesting sense. For syntactic categories, the idea predicts that surface order will convey topic/ focus/ presuppositionality by initial prominence, and secondarily special inflectional markings.

To some extent all this is testable, which is why I am writing this. First it suggests that in English, sentences that violate /presuppositional/ order will be more complex. Aside from the Agent > patient ordering constraint, very little research bears on hierarchies->order principles. A few studies i.e., (Gibson, E. 1998), involve the idea that "I" and then "you/s/he" are presuppositionally built into sentences, hence reducing complexity. Other studies (Frazier, L., and Clifton, Jr., C. 1996), suggest that mapping pronouns onto discourse-presupposed elements is actually easier than mapping them onto sentential antecedents not yet in the discourse representation. So there are a number of things to look at in English:

e.g.,

are the examples (1, 2 a) easier than (1,2 b), and (1,2 c) easier than (1,2 d)

- (1) a) The boy hurt a girl**
- b) A boy hurt the girl**
- c) The girl was hurt by a boy**
- d) A girl was hurt by the boy**

(these have to be tested with and without contexts because the use of "the" without a context is itself a bit odd).

- a) A/The boy crushed a/the rock**
- b) A/the rock crushed a/the boy**
- c) a/The boy was crushed by a/the rock**
- d) A/The rock was crushed by a/the boy**

Note that much prior research would suggest that (1, 2 d) should always be EASIER than (1,2 c), since (1,2 d) violates the NVN strategy, and there is some research to that effect. However, for purposes of testing the presuppositional hierarchy it should be tested with the indefinite article which has almost never been done. Testing it with the definite article may obliterate presuppositionality at least in English, because the article is so pronounced. Furthermore, what is really at issue is the difference between (1,2 a) and (1,2 c) vs that between (1,2 b) and (1,2 d). Finally, no one to my knowledge has contrasted sentences like (1,2 a) vs (1,2 b).

It is easy to generalize the paradigm to the full hierarchy e.g. (3a,b, c)

- 3a) You hit him, vs. He hit you**
- 3b) I hit you vs. you hit me etc.**
- 3c) The nurse loves the infant vs The infant loves the nurse**

(perhaps following Navajo, the differences may depend on being in the "inverse" construction, ie passive).

That is the first kind of thing to look at in English (or Korean, Russian, Chinese or Spanish). The other aspect of the model involves acquisition of the constraints as perceptual strategies. Again, little has been done on this directly. But it is sure easy to test for those who like to have kids act out short sentences. If my hunch is correct, four year olds are at a choice point in their treatment of agency and presuppositionality hierarchies: they can link them, so that the first item is generally taken to be the agent (as in English), while allowing odd surface orders when they are marked as in the passive: or they can impose the expected order as a surface order, regardless of the agency and use special markers when the agency is out of register with the order. The choice seems to be whether the hierarchy attaches to agency itself or to surface sequence constraints.

A conservative formalist might argue that this choice is a parameter that the child must learn: if parameter setting is a useful model of acquisition, so be it (but vide infra). What I am concerned to show is that the substantive content of the

constraint is possibly conditioned by the way language functions for the young language learning child. If this is supported experimentally, it will also support the interpretation of the hierarchy of order constraints as indeed due to how language functions. That is, it will stand as a clear case in which we can explain an apparently formal linguistic property as truly caused by how language behavior is assembled during cognitive development.

3.2 Up a tree! Linking elements in embedded and dominant fields

Movement has been a feature of transformational grammars since Harris. Movement occurs in the service of linking sites that a particular constituent has sequentially occupied during a derivation. It is striking that when only one of the exemplars of the constituent is pronounced (the typical case) it is the least embedded, the “highest” available in the constituent tree. This is the case regardless of whether the movement itself is conceived as “up” the tree (by far the most common architecture in the last fifty years), or “down” the tree (Richards, 1997) (Philips, 1996.)

Why? One can certainly imagine computational derivations in which the more embedded position is the pronounced one, which would have the same expressive power. Why does it turn out that the most salient site for a constituent that has one is the least embedded? Is there a functional explanation?

My answer to this may require some initial suspension of disbelief typical of drama in general. In this case, I consider the neurological foundations of humans’ ability to discover or create language. It is well known that the brain areas devoted to language in humans evolved as visual areas, as reflected in the localization characteristics of our simian near relatives. That is, the computational properties of the “language” areas in the human brain were shaped originally for the needs of vision in apes.

Suppose that the existence of movement as a mechanism within syntax recruits the mechanisms of vision originally related to the perception of *real* movement. Such co-opting of old functions for new ones is the watchword of evolution, and given the areas that language took over, it is not a wild speculation, only an adventurous one. We can make it less adventurous by actually exploring the question of whether real movement between sites in a visual scene makes the less embedded site more salient.

In studies of the perception of actual movement, this question translates into whether the movement is easier to perceive when it ends up in the less embedded site, compared with the reverse. Our experimental studies of this suggest that the answer is yes, movement is easier to perceive from more to less embedded parts of visual arrays (Lachter,, Weidenbacher. and Bever,. In preparation). We test this by presenting people with stimuli spaced so that they would ordinarily create illusory movement as in the classic gestalt phi phenomenon (a typical example is the perception of movement between alternating lights at a railroad crossing). We have

found that perceived movement out from a square within a larger square is easier to perceive than movement in the opposite direction. I won't give all the appropriate details and controls here, except to note that it is not the case that static visual stimuli are more salient in less embedded parts of a scene: rather, the relative salience of the less embedded position depends on there having been movement to it from a more embedded position. Real movement is easier to perceive visually from more to less embedded parts of a representation. If language has co-opted computational mechanisms originally adapted to the perception of vision, then they may determine aspects of linguistic computations that are otherwise arbitrary from the standpoint of syntax. The result is that syntactic movement, like that of a cat, is up the tree, not down the tree.

A similar question arises as in the causation of order hierarchies: is the salience of the less embedded exemplar of a constituent due to principles of the behavioral organization of movement, or is that a more general mental principle which lends salience to all phenomena that are less embedded? Most specifically, is it the result of why hierarchies and movement exist at all, as suggested by Kayne (1994) and Moro (1997). In essence, their proposals are that the role of linguistic movement is to create asymmetry in otherwise symmetric strings, to insure a specific manifest serial order that reflects an hierarchical asymmetry: that is, in a binary sequence [a,b], there is no structural basis for order, but if one of them is moved to an immediately higher tree, then a mapping rule which gives precedence to the less embedded member of a tree, will result in a specific output order, thus guaranteeing recoverability of the underlying relations. On this view, it is the computational properties of language which dictate the salience of less embedded constituents.

As in other cases, it is not easy to find a direct test to differentiate these hypotheses. In this case, we might turn to animals, since they can learn serial order, and study whether the order becomes fixed when a serial subset is treated as a constituent. That is, suppose we train an animal that, a-b or b-a are acceptable orders of a and b, and x-y or y-x are acceptable orders of y and x. (If it helps the reader envision how this would be done, imagine a monkey trained to point to different points on a compass rose, in a particular sequence) . At this point all the animal has learned is that a,b and x,y are cooccurrence pairs. Now, suppose we embed x,y within a, b, so the possible sequences are:

a, x, y, b
a, y, x, b
b, x, y, a
b, y, x, a

The critical question is: does the animal pick a particular order of x,y, now that it is embedded within a,b and hence a "constituent". Of course, this simultaneously tests whether the animal can learn structurally dependent contingencies at all (some recent studies suggest that it can), and whether serial order becomes fixed as soon as a sequence becomes recognizable as a constituent. With Fraser Wilson and Nick

LaMendola, I hope to be carrying out such studies with monkeys, and we shall see. A positive outcome will tend to argue that the order constraint results from an hierarchical structure, be it in language, perception or other behaviors. We shall see.

3.3 Aha! The eccentric mapping of meaning and form makes grammar fun to learn – or is it the other way around?

One purpose of language is to construct and convey meaning. In that case, why is the mapping between syntax and semantics so complex? Would it not be most functional if categories and units of meaning mapped univocally onto corresponding categories and units of form? Why does form jumble up the surface presentation of meaning, so that it is often impossible to extract it, except at the level of sentence, or even sometimes, discourse?

Following Chomsky, and Miller, (1963). „, we can view the problem of the meaning-form mapping as a problem of mapping between a hierarchical unordered multidimensional structure and a linear ordered sequence. The representation of meaning contains complex information, which has to be condensed into a linear string of symbols. Recursion virtually guarantees that the linear string will contain elements ‘related at a distance’ as opposed to combined or adjacent. Thus, there must be some mismatch between the linear groupings available in sentences and the semantic structures they express. On this view, much of syntax lays down restrictions on distantly related elements so that the relevant meaning can be recovered.

Thus, the first answer to the meaning/form mismatch is based on the incommensurability of the metalanguage of the two domains. Put in rampant functionalist terms, the purpose of syntax is to constrain serial and lexicalized forms so that they are reliably related to meanings. On this view, the answer to the semantic/syntax mismatch conundrum is that it cannot be avoided.

But clearly it *can* be avoided for isolated phenomena. For example, there is no channel or communicative reason why semantic gender could not reliably predict syntactic gender in those languages that have it. Yet this is rarely the case, if ever. Characteristically, syntactic gender must be learned as an arbitrary feature of each noun – there may seem to be some vague semantic field associated with limited subsets of a syntactic gender, but nothing systematic and always replete with exceptions. The corresponding phenomenon exists for verb conjugation classes: *sometimes* they can be identified phonologically, but rarely if ever, semantically. These are clear cases where nothing would be lost in expressive power if there were strict regularity between semantic and syntactic categories: so why aren’t there?

An operational way of putting this is to ask: why don’t language-learning children re-regularize the semantic/syntactic category lineup, disregarding their parent’s perversity? Of course, for a while they *do*, that is, there are reports from various

languages that young children sometimes pick one syntactic gender as the default for a new word, but also sometimes assign its syntactic gender based on some theory about its meaning. But that does not stick.

The language learning process may be the culprit. We all know and regularly genuflect to the concept that *language is innate*: that is, the child must contribute many – perhaps all – of the interesting structural properties inherent to the semantic/syntax mapping, because by virtue of the reasoning above, the syntactic universals are not transparent in the language experiences of the child. All well and good: the child has the innate capacity to invent grammar.

But where is the *motive*? It won't do to sit back and assert that the 'communicative urge' is sufficient: children communicate very well with ungrammatical utterances, indeed often without much in the way of utterances at all. A Markovian 8th order of approximation to English would produce and process most any sentence the child might need. Why does the child work and work away, to arrive at the kind of abstract grammar that children learn? Look at it from the child's standpoint – why does s/he work so hard to get it exactly right?

There probably are convergent reasons, including social identification, communication, the need to feel superior (over littler kids), among others. But I think there is another overriding internal motive: *language is a puzzle that is fun to solve*. More than any other indigenous skill, it affords one "aha" discovery after another, over an extended period, at least 8 years in my estimation. What other activity can beat that?

This argument presupposes several claims:

- a) people like to solve problems of a particular kind.
- b) Language presents that kind of problem.

We know that people like to solve puzzles, but what constitutes a really satisfying problem? It is hard to discover that by looking at real problems, even those instantiated in games, because real problems bring their own constraints and goals to the situation, thus muddying any analysis. A different way to approach the issue is to examine what makes an aesthetically satisfying experience. Aesthetic preferences – by definition in some cases – are free from any external goals, except the one of providing enjoyment. In this, we follow the original reason to study aesthetics voiced by the founder of such investigations – Gustav Fechner: such study reveals how the mind works in free action (*im Leerlauf*).

The fine arts are complicated because we pay for them, so it is more helpful to examine "folk aesthetics", entities that people prefer naturally. Consider for example, the golden section rectangle, which has the ratio of 1.62+ to one. Fechner documented that it is preferred in many ways, despite the fact that from the standpoint of obvious material functions (e.g., in construction blocks, or windows) it

is an infelicitous ratio. The explanation lays in a convergence of both general aesthetic theories, framed originally by Aristotle:

- a) **The Goldilocks Principle:** objects must have just the right amount of complexity, not too much, not too little.
- b) **Conflict resolution:** an object must stimulate some kind of representational conflict, and afford a resolution of it by accessing a new integrative concept.

The golden mean ratio affords an interesting kind of complexity, being the mathematical limit to both geometric and linear progressions. In perceptual terms, this means that if a square of the short side of the rectangle is subtracted from its area, what remains is another rectangle with exactly the same ratio (rotated 90 degrees). There is considerable evidence that the perception of simple figures (and perhaps everything) involves analysis into simple canonical figures such as squares. Thus, the interesting complexity of the golden mean rectangle is that on the one hand it can be analyzed by successive subtraction of squares, always leaving the same figure as residue, but the analysis is never entirely complete. The claim that this is an aesthetically optimal level of processing complexity requires independent verification, so at the moment it is a reasonable speculation only.

We have, however, investigated some implications of the claim that the golden mean rectangle affords a resolvable representational conflict. Consider first the square-subtraction method for a rectangle with a ratio of 1.5 to one: The two squares overlap to create three vertical rectangles of .5 to one: thus, the figure can be “resolved” in two dimensions by the propagation of a single figure (itself resolvable into two canonical squares of .5). Most important it does not matter for this analysis, which square is chosen ‘first’, the overlap collapses into a simple resolution either way.

The golden mean rectangle does not afford such a resolution: overlapping squares on the short side define a rectangle of .42 to one, which does not afford any direct resolution of the entire figure into a single two-dimensional figure. This sets up a conflict – should the left or right square be the one subtracted? This sets up the representational conflict. The conflict can be resolved by accessing a different level of representation, namely by analyzing the two squares as overlapping in three dimensions. Thus, on this interpretation, the golden mean rectangle stimulates a situation that meets Aristotle’s conflict-resolution model of aesthetic objects. It stimulates a conflict that is resolvable by accessing a different level of representational analysis.

This makes a neat prediction: the golden mean rectangle should stimulate more perception of depth (at least in its center) than other rectangle shapes. First, we noted that various schools of art that allow variation in frame shape, tend to converge on the golden mean rectangle for landscapes, compared with interior scenes: perhaps artists already know that the golden mean rectangle enhances depth. We have also studied the effect of frame shape on depth perception

experimentally. We use standard depth illusions, and find that they are enhanced within the golden mean frame, compared with other shapes (Nicholas and Bever, ms). So, there is one experimental confirmation of the general theory of aesthetics, at least in regard to conflict resolution. It also accords with intuitions about many experiences of explicit problem solving. In a classic case study of “insight”, elucidated by the Gestalt Psychologist Wertheimer, people solve the problem of how to x-ray tumors without killing healthy tissue, with the following stages:

- a) the x ray gun will kill the healthy tissue if it is strong enough to kill the tumor
- b) weakening the gun will allow the tumor to survive
(conflict)
- c) AHA, point the weakened gun at the tumor from different points each time and accumulate radiation of it, without killing any particular healthy tissue.

That is, accessing the idea of pointing the gun from different places (a new dimension) resolves the conflict. People just like to do that. Perhaps, conflict resolution by accessing unexpected levels of representation is innate.

To return to language in particular. We can now interpret the child’s tolerance of the persistent mismatch between semantic and syntactic units and categories: a language with such mismatches presents an ongoing series of problems to solve. The problems are *fun* to solve just because of the conflict between the different apparent representations – surface form and meaning units. Abstract syntax is the level of computation, which solves the constantly appearing conflicts between meaning and surface form. So, the function of the mismatch is that it makes multileveled grammar fun to learn, and therefore what is learned.

But now we can turn the argument on its head. Return to Chomsky and Miller’s observation that the problem that language sets is mapping a complex hierarchical semantic representation onto a serially ordered string. This necessarily involves a “compression” of the hierarchical information, (or decompression of the serial information), thus guaranteeing some kind of multi-staged derivation when mapping one kind of representation onto the other. The arguments I made concerning the child’s intrinsic motivational structure to learn abstract systems, requires the language learning child to have “fun” learning the kind of derivational sequences that languages must have. Result? In order for languages to be learned, humans must find multi-leveled structures fun to learn. Result of that? Humans find many problems and objects more satisfying when they exhibit multileveled properties. Thus, on this view, the characteristic human problem solving and aesthetic preference systems indeed have strong formal relations to language learning, but the causal direction is from language to cognition, rather than the reverse. This may turn out to be an unexpected kind of Whorfian result: a profound property of human cognition may be the result of a mental proclivity whose initial role is to guarantee the learnability of language in particular.

On either causal view, we can explain why the semantic/syntax mapping is irregular, even in those cases where making it regular would not impair the expressive power of language. Children are pattern learners, with ever more evidence that this starts virtually at birth (e.g., Gomez and Gerken 1999; Saffran, Aslin, and Newport, 1996). Suppose that nominal semantic and syntactic gender, or lexical verb semantics and syntactic subcategorization were in perfect register. The child would arrive at an easy pattern, tending to confirm the idea that semantics and syntax have simple mappings onto each other. Yet, that is critically false when it comes to non-lexical compositional forms. The child could be prematurely misled into an acquisitional cul-de-sac, expecting all semantic/syntax mapping to be similarly transparent. On this view, the existence of exceptions is critical to stimulating the right kind of learning.

This hypothesis, too, is subject to some testing, perhaps using “artificial grammar” or pattern learning in infants and adults. The theoretical question is easy, if the experimental design is hard: do learners actually learn structurally multileveled systems faster when the data are judiciously sprinkled with exceptions to their general rules?

4.0 Conclusion. What is Functionalist Investigation For?

I have tried to demonstrate how far reaching functional explanations of linguistic structures might be. Of course, showing that a linguistic universal might be explained by an extra-linguistic function, does not mean that it is not embedded as part of the innate structures that directly prefigure language. But it does open the question. Thus, the priority of canonical deep and surface word orders may not be innate, but rather the result of the early stages of language use by children; the fact that less embedded positions are more salient, does not mean that this is part of the formalo innate structures for language: rather the biological substrate for language forces links between ever more salient sites to work upward and not downward. Finally, the mismatch between semantic and syntactic categories may not be linguistically innate, but an historically winnowed properties of extant grammars, because they were the ones children want to learn for general reasons.

Why should a linguist care about these abstruse potential extra-linguistic sources for linguistic structures? I cannot dictate an answer, but I do know that the truth shall set you free. In this case, free from being required to embed every linguistic universal theoretically as derived from a linguistic structure. For example, the proposal that canonical word order is a universal linguistic constraint within a stochastic ordering of universal constraints, presupposes and supports a particular architecture for syntax – stochastic optimality theory. Indeed, the apparent merging of grammaticized requirement and statistical preference for canonical orders is a strong argument in favor of the architecture that allows such merging. But if canonical construction forms are themselves learned from the surrounding language, and perhaps shaped by the order of thought, then the claim that the

constraint is embedded as a central linguistic universal is gratuitous, and potentially quite wrong.

In times of yore in language science (when Eloise was starting out her training), there was disagreement about the ontology of grammars: on the one hand, ‘hocus pocus’ linguists argued that grammars are merely efficient organizations of linguistic data; ‘god’s truth’ linguists believed that there is a correct grammar. Most today are probably in the latter camp. But if getting it right only involves the correct description of an abstraction, it would matter only to those who hold linguistics to be a branch of mathematics: that is, ‘hocus pocus’ and ‘god’s truth’ linguists might end up agreeing that whoever is right, there are no implications for human biology or cognition.

Today, the biological basis for language looms gigantically in the background of linguistic investigations. None of us knows where linguistic science is headed, but many suspect that it will include considerations of the specific genetic foundations of language capacity: Indeed, speculations about the relation between linguistic universals and genetic substrate are finding respectability – premature in my opinion, but the goal is noble (Baker, 2001, Jackendoff, 2002). This makes it important in the concrete, to ascribe linguistic universals to their actual cause: we do not want to be looking for the genetic source for extra-linguistic features of the mind, when we focus on the unique genetic source for language. Conversely, we do not want to be looking for the genetic source for general cognitive features of the mind, if they are actually extensions of linguistic capacities. Put more softly, we want to understand the potential relations between linguistic and general cognitive properties, when exploring the biological and ontogenetic bases for either one.

REFERENCES

- Bever, T. G. 1970. The cognitive basis for linguistic structures. In R. Hayes (Ed.), *Cognition and language development* (pp. 277-360). New York: Wiley & Sons, Inc.
- Bloom, Lois. 1970. *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.
- Bresnan, Joan. 2001a. Explaining morphosyntactic competition, in Baltin and Collins (eds), pp. 11-44.
- Bresnan, Joan, 2001b. The emergence of the unmarked pronoun, in Legendre et al. (eds.), pp.113-142.
- Bresnan, Joan, in press. The lexicon in optimality theory, in S. Stevenson and P. Merlo (eds.), *The lexical basis of sentence processing: Formal, computational, and experimental issues*, John Benjamins, Amsterdam.
- Bresnan, Joan and Aissen, Judith. To appear. Optimality and functionality: Objections and refutations. *Natural Language & Linguistic Theory* 20.1.
- Bruner, Jerome. In the Annual Report of the Center for Cognitive Studies, Harvard, 1964.
- Chomsky, N., and Miller, G. A. 1963. Introduction to the formal analysis of natural languages. In D. Luce, R. Bush, and E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 269-321). New York: Wiley
- Frazier, L., and Clifton, Jr., C. 1996. *Construal*. Cambridge, MA: MIT Press.
- Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1-76.
- Gomez, R. L., and Gerken, L. 1999 Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70, 109-135.
- Jackendoff, R. 2002. *Foundations of Language*. Oxford, England: Oxford University Press
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass: MIT Press.
- Lachter, J., Weidenbacher, H., and Bever, T. G. In preparation

Limber, 1995. Personal communication.

Moro, Andrea. 1997. *The raising of predicates*. Cambridge: Cambridge University Press.

Newmeyer, Frederick, 1998. *Language Form and Language Function*. Cambridge, MA: MIT Press.

Nemeyer, Frederick. 2002a. Optimality and Functionality: A Critique of Functionally-Based Optimality-Theoretic Syntax. *Natural Language and Linguistic Theory*, 20, 43-80.

Newmeyer, Frederick. 2002b. A Rejoinder to Bresnan and Aissen. *Natural Language and Linguistic Theory*, 20, 97-99.

Nicholas and Bever – The Perceptual Basis for the Golden Mean, ms.

Philips, Colin. 1996. Order and structure. Doctoral dissertation, MIT, Cambridge, Mass.

Richards, Norvin. 1997. What moves where when in which language? Doctoral dissertation, MIT, Cambridge, MA.

Slobin, D. I., and Bever, T. G. 1982. Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, 12, pp. 229-265.

Townsend, D. J., and Bever, T. G. 2001. *Sentence Comprehension*, MIT Press

Saffran, J. R., Aslin, R. N., and Newport, E. L. 1996. Statistical learning by 8 month old infants. *Science*, 274, 1926-1928.

Willie, MaryAnn and Jelinek, Eloise. Manuscript.

Willie, MaryAnn and Jelinek, Eloise. 2000. Navajo As A Discourse Configurational Language. In Fernald, Theodore, B., and Platero, Paul, R., (eds). *The Athabaskan Languages. Perspectives on a Native American Family*. Oxford University Press.