22

The biolinguistics of language universals: The next years

THOMAS G. BEVER

22.1 Introduction

Bever (1970) outlined a program of research on the study of language universals:

> The concept of "language" is like those of "species" or "organ" as they are used in biological science … grammatical structure "is" the language only given the child's intellectual environment … and the processes of physiological and cognitive development … [O]ur first task in the study of a particular [linguistic] structure implicit in adult language behavior is to ascertain its source rather than immediately assuming that it is *grammatically* relevant … Many [an aspect] of adult … linguistic structure is itself partially determined by the learning and behavioral processes that are involved in acquiring and implementing that structure … Thus, some formally possible linguistic structures will never appear in any language because no child can use [or learn] them.   (Bever 1970: 279–80)

The first sentence of this essay expressed the then contemporary biolinguistic approach to language. The later sentences began a program of research and theorizing aimed at exploring the limits of the biolinguistic dogma for language structure as it appeared to many in 1970. In brief, it was:
Language capacity is innate, with the following supportive considerations:

(a) The poverty of the stimulus: the child does not receive or try out an adequate number of sentences to guide any inductive learning system.
(b) The structures that the child easily imputes to language are too abstract to be explained without an internal creative mental ability: e.g., derivations and even phrase structure are themselves abstract structures that organize established lexical representations.

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/13   Time:09:50:39
Filepath:d:/womat-filecopy/0001977625.3D386

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

386   *Thomas G. Bever*

(c) Acquisition appears to occur in big steps without special motivation for each step.

(d) There is a critical period for learning a first or second language, which is consistent with endogenous maturationally timed behavioral abilities in animals, such as song in certain bird species.

(e) Syntax and other linguistic components are self-coherent mental modules, suggesting a unique neurological basis especially for syntax.

(f) The specialization of the left hemisphere for language in humans is unique, and may be uniquely related to the language ability itself. As a corollary,

(g) There is one normal neurological organization for language in humans.

The research program since CBLS has been haphazardly devoted to modifying these classic biolinguistic assumptions about language, in the context of its acquisition, use, and representation. The general goal has been to explore the possibility that apparent language-specific universals in fact derive from more general cognitive and neurological systems. A series of investigations have suggested that the following points are worth considering, if not proven.

22.2 The interaction of statistical (inductive) and categorical (deductive) processes

The most enduring thread of thought in all these developments has concerned the role of induction and statistical knowledge in structural language behavior and acquisition. In each case, demonstrations and arguments have shown that language behavior involves a critical interplay between statistical strategies and structural processes. This started with a discovery outside of language, that two-year-old children actually conserve quantities under transformation that make them look bigger or smaller. At the time, this was considered important because the Piagetian findings had started with four-year-old children, whose quantity judgments are basically dependent on the apparent size of a display. The question then was: why do four-year-old children make worse judgments than two-year-olds? Our answer was that four-year-olds have mastered a true statistical generalization which then guides their judgment: *things that look bigger have more "stuff" in them* (Mehler and Bever 1967; Bever, Mehler, and Epstein 1968). Hence, the discovery of a U-shaped cognitive developmental curve: two-year-olds tend to be correct in many simple cases, where four-year-olds become incorrect, and six-year-olds are correct again. This developmental phenomenon in general cognition

Biolinguistics of language universal

generated a number of investigations of different kinds of cognitive capacities (Bever 1962a, 1962b).

Early stages of language acquisition followed this pattern, as initially presented in CBLS. The two-year-old understands above chance levels sentences that violate the canonical NVN (agent–verb–patient) pattern of English, but the four-year-old is much more dependent on the canonical order, even when it leads to incorrect reversed interpretations. The six-year-old recovers much of the correct initial interpretive skill.

Adult sentence behavior also depends critically on statistical processes: "whatever the limits of associations, habits exist and dominate most of life. Any adequate model of comprehension must find the appropriate computational locus for their operation and influence" (Bever and Townsend 2001: 149). The appropriate model turns out to be a rehabilitation of the classic analysis-by-synthesis model, a particular model in which:

1. Associative information operates on relatively superficial representations and is immediately available.
2. Readily available surface information includes the lexical items in sequence, a rough phrasing structure, and a likely meaning.
3. Syntax is derivational—it involves a series of steps in building up sentence structure which can obscure the initial computational stages in the surface form. (Bever and Townsend 2001: 50)

As we put it: we understand everything twice.

This perspective clarifies the negative and positive role of connectionist modeling, which still enjoys considerable popularity as the sole theoretical solvent for modeling psychological phenomena:

The negative . . . is that connectionist models can capture only habits and are therefore inadequate in principle to capture structural processes. The positive . . . is that connectionist models can capture only habits, and they are an important new tool in the study of how habitual knowledge interacts with structural processes. (Bever 1992: 213)

22.3 The psychological reality of grammar

Over sixty years of apparent dramatic changes in generative theory, a constant feature has been some form of derivation relating an inner to an outer form of sentences (for history, see Bever 1988; Townsend and Bever 2001: chapter 3). Thus, the biggest ongoing puzzle presented to psycholinguists concerned with the role of syntax in adult language behavior is the following conundrum:

> Sentences are externally serial, (i.e., "horizontal"): derivations are internally hierarchical, (i.e., "vertical").

That is, the computational domain of a derivation can embrace entire clauses and sentences, while the immediate processing appears to be one word after another. The analysis-by-synthesis model described above integrates the serial and derivational aspect of sentence processing, but in a rather brutal manner—that is, it simply joins together the two kinds of representational systems. So, it is critical to gather evidence that the derivational component is evident during language behavior. At one point, this goal was expressed in demonstrating the role of the "deep" structure of sentences and transformations (e.g., Miller 1962; Mehler 1963; Clifton and Odom 1966). But as syntactic theory has evolved, the notion of what counts as "deep structure" has evaporated and become more continuously composed, not presenting a single unified structure in one place (Chomsky 1995; see Boeckx 2006 and others). What has endured in various forms is the notion of abstract empty categories that fill phrase structure positions as needed to maintain consistent generalizations about structure. For example, raising constructions ('John seemed [e] happy'), passives ('John was hit [e]'), unaccusatives ('John tripped [e]'), and others involve positing an unpronounced copy or placeholder of the overtly pronounced noun. This kind of rather bold claim afforded a series of moderately successful early investigations of the presence of such empty entities in English (Bever and McElree 1988; McElree and Bever 1989; MacDonald 1989) and Spanish (Bever and Sanz 1997; Sanz et al. 1992). The importance of these "noun phrase" traces is that most other syntactic theories do not have formal equivalents (unlike *wh*-trace). Thus, experimental demonstration of the presence of such traces during comprehension not only gave special behavioral support to the derivational assumptions of generative grammar, it motivates the assignment of derivations in a model of language behavior.

22.4 The computational basis of modularity of language

An important claim about syntax and its uniqueness rests on the idea that it is a module, architecturally distinct from other kinds of knowledge, both in its physical instantiation in the brain and its computational structure. This idea was complemented by a consideration of why language and many other cognitive skills appear to be architectural modules, held apart in how they are represented (Fodor 1983 crystallized this, but see also Forster 1979). Bever

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/2013   Time:09:50:39
Filepath:d:/womat-filecopy/0001977625.3D389

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

(1992) argued that many, if not all, aspects of "modularity" are based in computationally immiscible systems, not architecture:

[There is] a point of logic underlying the necessity of modularity when different kinds of representational systems are concerned. If the computational languages of two systems differ, one cannot affect the internal operations of the other. This does not necessarily demonstrate an architectural boundary between them, because their mutual computational opacity would lead to such discontinuities of influence anyway.   (1992: 183)

A further argument that modifies the notion of modularity derives from the isolation of several dimensions along which cognitive skills can differ. The first of these is the differentiation of left- from right-hemisphere processing in a range of areas, not just language; the second is gender, showing that female humans (and rats) process spatial knowledge differently; the final one, discussed below, is the different neurological representations of language in people with vs people without familial left-handedness. In brief,

I am arguing that there are regularities in general cognition independent of modality, and that those regularities are genetically coded. I have isolated three cognitive dimensions along which biologically based populations differ...there still may be specific neurological mechanisms that are responsible for recruiting general cognitive capacities in ways specific to each modality.   (Bever 1992: 179–80)

22.5 Neurological organization of language and its variants

Part of the basis to the claim that language is innate is the extent to which it has a unique dependence on a particular neurological organization. But aside from the logical necessity of linguistic modularity, how distinct is the neurological organization for language from other natural and special skills? The simplest view would be that there is a particular computational subregion of the brain that is unique to humans, and hence the neurologically evolved engine for language. In the 1950s to 1970s, the standard view on all this was that each hemisphere of the human brain has dominance for particular modalities, the left for language and logic, the right for space and artistic activities (e.g., Kimura 1961). This was in contrast with a much older view that there is a computational difference between the hemispheres, the left being structural or "propositional," the right being "associative" (Hughlings Jackson 1878, 1879). Music was a critical case: on the one hand it is serial and offers a form of hierarchical structure, on the other hand it is an art form. To this end, dichotic listening studies had demonstrated that melody recognition was better in the left than the right ear (e.g., Kimura 1964). The case seemed closed

until there were studies that differentiated *how* the listener cognized music. Separate studies showed that musically experienced amateurs apply a kind of "chunking" of melodies, similar to the kind of phrase segregation in language while musically disengaged people do not (e.g., Tan et al. 1981) This differentiation also made all the difference in which hemisphere is dominant for melodies: musically disengaged subjects indeed show better recognition in the left ear; but musically experienced amateurs show better recognition in the right ear (Bever and Chiarello 1972). This result played a role in resuscitating the more traditional idea, recast as the left-hemisphere processing in an "analytic" or hierarchical mode, while the right hemisphere is characteristically "wholistic" or associative. Thus, the priority of the left hemisphere for language is the result of its general ability for more complex computations and representations, not unique to language (Bever 1975). This could result from a very small computational difference at birth:

Suppose the only difference between the hemispheres is that the left hemisphere is more capable during early childhood. That is, suppose that the two hemispheres function identically at birth but that the processing capacity of the left hemisphere is larger. This substantive claim, together with several other independently justified premises is sufficient to account for the early appearance of cerebral asymmetries and their continuous development. (Bever 1980: 212)

Various theories have followed as to *how* it is that the left hemisphere is more powerful computationally; for example, a recent claim is that the left hemisphere has a dominant fast circuit, roughly the rate of phonemic streaming, while the right hemisphere has a dominant slow circuit, the rate of syllables (Poeppel 2003; Giraud et al. 2007).

Infant data is prima facie inconsistent with the idea that language asymmetries flow from a general computational difference between the hemispheres. First, Mehler and colleagues have shown that neonates respond to forward speech with more activity in the left hemisphere than to backward speech (Peña et al. 2003); furthermore, Dehaene-Lambertz et al. (2010) show that speech and music activate different left-hemisphere areas in two-month-olds. These results suggest a strong innate bias in which the newborn left hemisphere is "tuned" to speech, with specific neurological areas differentially involved. On a strong version of this view, the early acoustic bias may "drag" language-specific phonetic, syntactic, and semantic aspects of language into the maturing left hemisphere, without the influence of any other special differences between the hemispheres: more general adult differences between the hemispheres would be parasitic on this early bias. So, we have a conundrum: there is a specific neonatal tuning of the left hemisphere for speech, but there is a modality-general computational and speed superiority in the adult.

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/13   Time:09:50:39
Filepath:d:/womat-filecopy/0001977625.3D391

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

It will take more research to sort out which is causally efficient. An evolutionary interpretation could be that the neonatal tuning to speech input evolved as a way to guarantee that the more powerful hemisphere would be the one that ends up computing language—arguably the most complex of all natural human behaviors.

Our recent studies may lead to a research tool to study this question in the infant as his or her neurological and linguistic maturation emerges: we are showing that asymmetric aspects of language representation may be normally quite different as a function of genetic background of asymmetries. In particular, right-handers with a left-handed familial background have characteristically different behavioral responses to language and other tasks, from right-handers without familial sinistrality. This finding opens up a new line of genetic investigation into language, independent of a genetic anomaly: it also gives further support to the idea that language structure is not itself caused by gross aspects of neurological structures: rather it finds the best available representation for particular brains. The implications of this are discussed below.

22.6 Language, structural capacities, and related phenomena in animals

If language has a biological origin, one expects that there will be analogues in animals that will reflect a common evolutionary background or genetic make-up. Our work in this domain started with relatively traditional methods—training pigeons to peck a series of four colors in a particular order with Skinnerian schedules of reinforcement. The hope was that eventually we could "chunk" separate sub-orders either by the method of training, or rhythmic spacing, so we could study the effect of embedding one sequence in another, recursive sequences, and so on. Of course, this was a very naïve and ambitious program: pigeons become quite adept at pecking a sequence of colors, but only after thousands of trials, and after careful building up of the sequence, first with doublets, then triplets, and finally all four. Careful analysis of their error and latency patterns suggested that they in fact did eventually rely on an internal representation of the sequence, but as a string without internal structure (Bever et al. 1980). Bever (1984) argued from these data that they support the notion of "representational abstractionism," a principle in opposition to the traditional "representational reductionism," that one should always assume the least powerful computational mechanism found in a behavior: rather, Bever suggested that it may be more fruitful to start with

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/13   Time:09:50:39
Filepath:d:/womat-filecopy/0001977625.3D392

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

the assumption that a behavior calls on the highest computational mechanism of which an animal is capable.

Of course, no one expected pigeons to talk, or peck out a version of Morse code. Greater hope was laid on sign language in chimpanzees. For about four years, a group at Columbia University attempted to give a chimpanzee a human-like environment, with intensive training in American Sign Language. (Bever's initial contribution was to name the chimp "Nom Chompsky," conventionally known as "Nim Chimpsky" by virtue of vowel raising.) At first, Nim appeared to develop some kind of sequential structure, modeled at least on a Markov system, with some loops and serial contingencies. However, as his utterances became longer, it was clear that he used some form of excitable word salad, partially imitating recent utterances by his trainers. As an experimental failure, this study remains one of the most definitive demonstrations of what a neighboring species can *not* do (Terrace and Bever 1976; Terrace et al. 1984).

The neurological organization of behavior in animals is very often studied to gain a better understanding of homologous structures in humans. This not only serves to provide broad biological background of brain structure; it can be important in the study of medically relevant issues. The first kind of study is behavioral, exploring peripheral asymmetries in animal perception that may be similar to humans'. First was a study of tone sequence discrimination in rats, with the tones presented to one ear or the other. For this, O'Connor et al. (1983) used a constructed rat listening station, in which rats were trained to position themselves so that each ear flap covered its own adjacent air-earphone. The paradigm was a go/no-go one, in which the rat pushed a lever with its nose to positive instances of the stimulus sequences. The results clearly showed with a number of rats a right-ear advantage, which increased as the target sequence of tones was longer. The similarity of this kind of result to human research also confirmed the view that cerebral asymmetries like those of humans are general in mammals.

Further perspective on cerebral asymmetries in humans comes from a clear but unpublished study of asymmetry for a symbol system in the dolphin. Herman et al. (1990) had discovered that once a dolphin was trained to respond to a "language" with a set of hand-signed commands, it would actually respond as well to a small TV presentation of the signals (shown through the side window of a tank). It turned out that presentations to the right eye (left hemisphere) of the dolphin resulted in faster and more accurate responses than to the left eye. Unfortunately, one dolphin was not sufficient for publication and the complexity of running more was too great to follow up (Morrel-Samuels et al. ms).

Of course, as in humans, the underlying mechanism of peripheral asymmetries is functional asymmetry of the hemispheres. The study of cerebral asymmetries in rats offered the possibility that asymmetries in humans are not unique, and also the possibility that rat asymmetries could be used to study normal and abnormal function of asymmetrically represented skills in general. To this end, we used mild spreading depression induced by administering potassium chloride to one hemisphere or the other, to "fuzz up" one hemisphere or the other, while rats learned particular patterns in a radial arm maze (which arms were regularly baited, which were not). The results showed that indeed, the left and right hemisphere of the rats learned the maze in very different ways—the left appeared to have a localist serial representation, while the right had a more global "world view" representation (Lamendola and Bever 1997). This established the potential for further study of what the best therapies might be for left-hemisphere-based aphasia in humans. It also indicates further that cerebral asymmetries of the general computational sort found in humans may be characteristic of other mammals.

22.7 Extralinguistic sources of language universals

These studies circumscribe some of the central and peripheral claims that have been associated with the innateness of the language capacity. But the central fact remains: certain essential features of language are genetically available to children. What those features are and how they interact with interface constraints, neurological capacities, and formal laws continue to be the central questions for research in the language sciences.

In a later discussion, Bever (1982c) explored the possibility that the essential structures of language ("narrow syntactic universals" in today's sense) are to some degree not caused by learning, but formal extrinsic patterning forces: in that paper, I claimed that we have confused the fact that language acquisition proves that something is innate with the claim that what is innate must be the underlying theory. I also argued that the structure of some linguistic universals is due to factors not uniquely intrinsic to humans. Regarding this issue, Bever (1982c) discussed two possibilities:

1. *The essence of language (today aka "narrow syntax") is a natural form (Platonism)*

This essay explores some implications for the study of language acquisition of the view that the essential formal characteristics of language are not human in origin . . . they are universal abstract objects whose properties are uncaused [by human cognition] . . . (Bever 1982c: 448)

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/13   Time:09:50:39
Filepath:d:/womat-filecopy/0001977625.3D394

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

394   *Thomas G. Bever*

a result of

2. *The essence of language is the result of physical law(s)*

...If the essence of language is the result of a physical law that becomes relevant only when there are complex living systems like human brains, then the essential nature of language would be literally a law of the universe, not a law of the human brain or human history. (Bever 1982c: 448)

Consider an example of how a physical law might emerge in human behavior but not be caused to have its essential properties by that behavior, like upright walking. We could stipulate as well that walking in humans is "innate": it is learned with little specific training and involves innate brain mechanisms. But one would not conclude that the physical laws with which the brain mechanisms interact are also caused by those mechanisms; rather, the physical laws are the result of the basic nature of matter, not the human brain.

...One could make the parallel argument about the physical basis of language...the human brain may be physiologically adapted to learn language but if language is the effect of physical laws, then the structure of the brain cannot explain why language is the way it is. (Bever 1982c: 448)

What are the implications of this for studying language learning and representation? As I discussed then, the essence of language is no longer necessarily ascribed to purposeful evolutionary causation. It could instead be the result of the emergence of sufficient (mental or physical) complexity for humans to become susceptible to the relevant forms of language.

This view would explain why...certain common cognitive processes never occur in language...such cognitive processes are not part of the extrinsically determined essence of language...if the essential features of language are real...then they are not caused by the mechanisms of human evolution or learning. (Bever 1982c: 440)

The implications of this view for acquisition of language are also explored in a series of earlier and later papers that emphasize the cognitively intrinsic motive individual children may have to acquire their surrounding language (Bever 1987, 2009). First, it was argued that accessing and internalizing a grammar plays a critical role in the acquisition of actual language behaviors, providing a cognitively consistent representation of the language across different language behaviors. For instance, in Bever (1975) it is stated that "The reason that a psychogrammar exists is because of the vital role it plays during language acquisition, much of which occurs during the first five years of life. The psychogrammar is needed during that period to mediate between the systems of speech production and perception. It is the internal translator that regulates conflicting capacities which arise as each of the two systems of speech behavior develops separately: if one system gets ahead of the other the psychogrammar can equilibrate their capacities" (Bever 1975: 74).

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/13   Time:09:50:39
Filepath:d:/womat-filecopy/0001977625.3D395

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

. . . I have sketched an argument about the role of a psychogrammar which draws on a rationalist view about the nature of knowledge and behavior. The general view is that the mind is composed of partially distinct systems, which interact with each other, relying in part on internal languages to translate from one capacity to another and to regulate differences in internal capacities that bear on the same class of external behaviors. A psychogrammar is an example of such an internal communication and regulating system. It regulates the relations between the emerging system of speech perception and the emerging system of speech production. Children need the psychogrammar for this purpose and they need to elaborate it as the other systems become more complex. Adults do not need it anymore, but we are stuck with it. It simply refuses to wither away. (1975: 73–4)

Bever (1981) used this interpretation of the early role of grammar in mediating emerging systems of language behavior, to explain the "critical period" for learning language: once the behavioral systems are brought largely into confluence, the grammar apparatus is no longer needed. This model draws on biological notions of critical periods in morphogenesis, based on "decoupling" of adjacent cell systems after they are basically oriented.

The main following point has been to show that language learning by the individual may be motivated at least in part by the intrinsic enjoyment of solving his or her language as a problem. That is, language learning is a special application of general principles of human problem solving and learning, activities that humans find especially exciting (eliciting the well-known "aha" reaction of excitement when a problem appears to be solved; Wertheimer 1945).

Suppose the child treats discovering the syntax of her language as one of the first big life problems to solve. This would explain it as motivated, not by the urge to communicate (as in the usual behaviorist explanation), nor as forced primarily by maturation (as in the traditional biolinguistic explanation), but as an activity that is cognitively intrinsically thrilling and implicit fun. That is, the child learns the language because it is an exciting, self-stimulating thing to do (Bever 1987, 2009). At the same time, current sociolinguistic research reminds us that language variation serves an important group-identifying purpose (see articles in Eckert and Rickford 2001). On this integrated view, children are determined to solve the problem of how their native language works because it helps them be "just like" the grown-ups around them: the cognitive thrill involved in successive solutions to how the adult system works provides stage-by-stage feedback and intrinsic reward. In this sense, it is not unlike the motives sometimes ascribed to why children learn to walk (based in part on the failure of such learning in feral children; e.g., McNeill et al. 1984).

Recently, we have formulated this in the framework of an analysis-by-synthesis model of language acquisition (Townsend and Bever 2001; Bever 2009). On this model, children apply both inductive and deductive

396   *Thomas G. Bever*

computations for hypothesis formulation and confirmation. The overall goal is to find a coherent structure for the language experiences that systematizes the relation amongst and between meanings and forms. This model makes several kinds of predictions:

(a) Languages should exhibit statistically valid patterns, independent from structural constraints. This is a necessity for the inductive component of the analysis-by-synthesis acquisition model to have data to formulate hypotheses for structural analysis based on the child's structural, deductive, language component. A simple example of this is the universality of a "Canonical Syntactic Form" in every language, a statistically dominant pattern across constructions. Recently, Bever suggested that the canonical form explains the linguistic stipulation of the "Extended Projection Principle" (EPP), a constraint on the surface form of sentences which is not motivated by any general syntactic principles (Lasnik 2001; McGinnis and Richards 2006). The universal necessity of such canonical forms serves as an example of the kind of explanation discussed in CBLS (Bever 2009): languages that do not have a canonical form are structurally possible, but would never be learned, and hence are never attested. Thus, we can explain ~~the~~ EPP as a function of language learning, not as a universal property of formal syntax.

phenomena

(b) In English, for example, it is critical that the canonical form both have a near universal surface appearance, but also have critical differences in some of the mappings of that surface form onto thematic relations. In English *almost* every sentence with the canonical surface form, assigns the initial noun phrase "agent" or "experiencer" status in relation to the following predicate (Bever 1970). But it is critical for the model that not *every* such sentence is mapped the same way. This variation sets a problem for the child to solve: what is the overall structure that accounts for both the surface features and the variation in the thematic mapping? This ~~calls on~~ application of the structural component of the dialectic involved in building up syntactic knowledge.

elicits

(c) The problem-solving model can mitigate the "poverty of the stimulus," by utilizing the canonical form to generate sets of meaning–form pairs that the child has not yet experienced. This helps the language-learning child to be a "little linguist" (Valian 1999) without having memorized a large number of form–meaning pairs and without querying the adult world the way grown-up linguists do. A classic reflection of this is in the research of Ruth Weir (1962), showing that children manifestly "practice" to themselves the paradigms in their language—most important is the apparent fact that they utter sentences in canonical frames that they have never heard.

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/13   Time:09:50:39
Filepath:d:/womat-filecopy/0001977625.3D397

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

(d) There are implications for language history and change. For example, Bever and Langendoen (1972) analyzed the ~~effects~~ consequences when the rich inflectional system of Old and Early Middle English collapsed, under the weight of its own complexity and the influx of Romance vocabulary: certain OE and ME sentences had unique inflections that disambiguated a main-clause postverbal noun as an object, blocking it as the subject of a following complement/relative clause. When the inflections were lost, the noun could now fit into the "NVN" pattern, and was subject to being incorrectly interpreted as the subject of the complement clause. At this point, the complementizer, which had been optional in such sentences, became obligatory. Bever and Langendoen argued that this kind of process is part of the dialectic between what makes a language easy to learn and easy to use, which is an important engine for language change.

22.8 Two topics for the future

Now consider a few topics that I think will be important areas of near future research.

22.8.1 *The real poverty of the stimulus*

One of the major enduring touchstones of first language learning that motivates nativist claims is the so-called "poverty of the stimulus." The child never hears or tries out enough sentences to account for the rapid acquisition of structural principles. This has been an important factor in arguing that children must have intrinsic knowledge of critically important linguistic universals—they do not have to learn them from the environment, rather they contribute them to their linguistic experience.

Examples of how the child experiences and assigns abstract structure-dependent analyses to word sequences generally assume that the child has achieved lexical analysis of his or her input, and faces only the problem of how to organize rules that govern explicit lexical sequences. But this oversimplifies the child's actual problem. First, even child-directed speech is often elliptical and ungrammatical. This problem is exacerbated further by a property of normal conversation that is usually ignored, but is beginning to receive scientific attention: in everyday speech, many acoustic details are slurred or even omitted. This can be demonstrated by showing that fragments several "words" long are impossible to recognize in isolation, but pop into complete clarity (for native speakers) when heard as part of an entire sentence (Pollack and Pickett 1964; Warner et al. 2009; Tucker and Warner 2010; Dilley and Pitt

2010; Gahl, Yao, and Johnson 2012). Consider first an approximate transcription of an example from adults talking to each other in a normal conversation (this is an actual example provided by N. Warner).[1]

(1)   —chlnnthu——

Try pronouncing this to yourself (hint: it is actually three words). Now look at a longer sequence in which the example was embedded:

(2)   An err we-er chlnnthu spah

When listeners hear the surrounding material, the excerpt immediately pops into consciousness and what one "hears" is:

(3)   And err we were chillin' in the spa.

Another example is:

(4)   Tyuv

(Hint: this is three words.) It is completely incomprehensible by itself, but when a latter portion of the longer sequence is included it is comprehensible:

(5)   Tyuv taimta toktme

Everyone immediately hears this as:

(6)   D'you have time to talk to me?

First, such facts demonstrate clearly that the minimal phonetic unit of comprehension is not the word; they demonstrate that comprehension must be operating with parallel hypotheses at several interactive levels—syntactic and phonetic computations proceed in parallel with frequent cross-checks at specific points. One can expect that where those cross-checks occur will be the focus of ongoing research, now that we have tools that can chop running speech into a full range of possible units. An initial hypothesis is the *phase*, the unit of syntactic structure that has just enough content for semantic analysis. (Phase theory is an active research area in linguistics, so the reader should be skeptical about details by the time this chapter is published, never mind a few years later. See Boeckx 2008 for a lucid explication of the technical issues.) Indeed, it is now an interesting research question whether phases are the "true" units of comprehension that the many "click" experiments attempted to define (Fodor and Bever 1965; Garret et al. 1965; Bever et al. 1969): stay tuned.

---

[1] The reader can hear examples like it on her website: <http://www.u.arizona.edu/~nwarner/reduction_examples.html>.

A related phenomenon is our conscious but apparently false perception that we understand the speech and hear it serially. For example, when the two conversational excerpt examples above are heard in their original context, the conscious phenomenology is that they are understood simultaneously with their serial order. This is especially surprising in 'dyuv,' where the critical (and incomprehensible) isolated sequence is *followed* by the crucial contextual material. The striking fact is that we are not aware that we could not have analyzed the initial sequence until the later material was heard: rather we are convinced that we understood it as it was phonetically presented. This simple fact demonstrates that language comprehension proceeds in sequences of "psychological moments" in which actual processing moves both forward and backwards. This phenomenon has barely been touched in the language sciences, but is clearly fascinating and will have profound implications for consciousness theories, once it is better understood (e.g., see Fraisse 1984 and prior references to the effect that psychophysically one can show that the "psychological moment" is about two seconds, roughly what is needed to account for the language phenomena).

Now consider the implications for the language-learning child. There is some evidence that "motherese" is somewhat clearer than normal conversations in many cases (Bernstein-Ratner 1996; Bernstein-Ratner and Rooney 2001), it may use devices to clarify word boundaries (e.g., Aslin et al. 1996) and it may be that infants prefer motherese when they have a choice (e.g., Fernald 1985; Cooper et al. 1997). But it is likely that the vast majority of speech that children hear is between adults or older children, and there are considerable cultural differences in whether motherese is used at all (Lieven 1994). Furthermore, various studies have shown that the syntactic or phonetic quality of the child's input may bear little relation to the child's emerging language (C. Chomsky 1986; McColgan 2011). In any event, well-articulated motherese may not be dominant even in child-directed speech. Consider a transcribed example from a real motherese sentence. First, attempt to understand the fragment below (five words!), taken from an actual utterance by a mother to her child:

(7)   Gtmnre'pm

Now see the whole utterance below. Try sounding out the phonetic version alone to see if you can (suddenly) understand the whole utterance. In the acoustic version, the final excerpt immediately pops into perfect comprehension, with the conscious intuition that the entire utterance was reasonably clearly pronounced.

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/13   Time:09:50:40
Filepath:d:/womat-filecopy/0001977625.3D400

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

(8)   Oh good, mumy pt thoz ma?zeenz weh so yu ca~t gtmnre'pm

(9)   Oh good, mummy put those magazines away so you can't get them and
       rip them

It is amazing enough that adults can understand conversational speech like
this. For a child the problem is doubly compounded, since its grammatical
knowledge is incomplete, and it has not yet had time to build up complex
language patterns. This simple fact vastly increases the poverty of the stimulus
problem, since in many cases the child may not be able to even encode the
utterance in enough detail to serve as a learning model. It is worth noting that
the example utterance is itself a coordination of two separate canonical
clauses (mommy put those magazines away, you can't get them and rip them).

There is an important implication of these analyses for the architecture of
Universal Grammar. Over many years, it has been argued that linguistic
processes are *structure-dependent* (Chomsky 1980). This is in contradistinc-
tion to the idea that Universal Grammar uses rules that are sensitive to serial
order. Rather, rules are characteristically sensitive to hierarchical structure.
This part of Universal Grammar has been shown to account for pathways to
language in first-language acquisition (e.g., Crain and Nakayama 1987 and
many others). Recent attempts have been made to show that serial learning
models can converge on such sensitivity (Perfors et al. 2006; Reali and Chris-
tansen 2005) but such models fail to generalize realistically (Berwick et al.
2011). Recently it has been shown that adults can learn serial rules but in so
doing they utilize different brain areas than characteristic of language (Musso
et al. 2003; Moro 2011). In the current "Minimalist" treatments of language,
hierarchical trees are constructed as sets, that is, without serial order con-
straints (Chomsky 2007b). On this view the surface order in language is
imposed by how it interphases with our systems of input and output, but
the actual computation of linguistic rules operates strictly on the hierarchical
structure and relations of units up and down that structure. The notion of the
time-free psychological moment in language understanding (and possibly
language production) is consistent with the computational irrelevance of
the overt language sequence: thus, the comprehension system is building
chunks of hierarchically organized structures which themselves may be
internally order-free, corresponding to order-free processing of the input.

22.8.2 *Normal genetically controlled variation in neurological*
representation of language

The usual reason for studying genetic effects on language is to demonstrate
evidence that language is "innate" in some interesting sense that differentiates

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/13   Time:09:50:40
Filepath:d:/womat-filecopy/0001977625.3D401

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

it from heritability of "general cognition," "communicative capacity," and so on. Thus, there are investigations of spared syntactic capacity in Williams syndrome children (Bellugi et al. 1994; Clahsen 1998; Zukowski 2004), as well as children with severe motor disabilities; conversely, there are forms of selective impairment of language in, e.g., Turners syndrome (Curtiss 2012) and "FoxP2" children (Vargha-Khadem et al. 2005). In each of these cases, the usual method (in principle) is to isolate a particular genetic abnormality and relate it to the selective sparing or selective impairment of language ability, thereby making more specific the claim that language is "innate."

In our behavioral research of many decades and recent neurolinguistic research, we have adopted a different method to provide converging information about the heritability of how language is used and represented. In particular, we have tracked the effects of familial left-handedness in right-handers. Many thousands of questionnaires have shown us that about 45 percent of all college students are right-handers with familial left-handedness, and an equal percentage of right-handers without familial left-handedness. Thus, we can use familial handedness as a tool to explore differences in how language is used and represented in two equally large "normal" populations.

Historically, Luria (1970) and colleagues (Hutton et al. 1977) noted that right-handed patients with left-handed relatives (FS+) recover faster from left-hemisphere aphasia, and show a higher incidence of right-hemisphere aphasia than those without familial left-handers (FS−). They speculated that FS+ right-handers have a genetic disposition towards bilateral representation for language, which often surfaces in their families as explicit left-handedness.

Over many years of behavioral research, we have found a consistent behavioral difference between the two familial groups in how language is processed, which may explain Luria's observation. Normal FS+ people comprehend language initially via individual words, while FS− people give greater attention to syntactic organization. A simple demonstration is that FS+ people read sentences faster and understand them better in a visual word-by-word paradigm than a clause-by-clause paradigm, whereas the opposite pattern occurs for FS− people. Another example is that if words in a short essay alternate in isolation between the ears at a normal rate, FS+ people understand the essay better than if the words are presented all monaurally; the converse is true for FS− people. Iverson and Bever interpreted this as a result of the relative segregation of each word from the adjacent ones in the alternating ear condition, making it easier for FS+ people to recognize each word separately (these studies and others are reported in Bever et al. 1987, 1989; Bever 1992). In another set of studies, Townsend and colleagues reported that

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/13   Time:09:50:40
Filepath:d:/womat-filecopy/0001977625.3D402

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

recognition of an auditory probe word from a just-heard sentence fragment is faster in FS+ people than FS− people, while the latter are more sensitive to the overall grammatical structure of the sentence fragment (main vs subordinate clause; Townsend, Carrithers, and Bever 2001).

The bilateral representation of language in FS+ people may be specific to lexical knowledge, since acquiring that is less demanding computationally than syntactic structures, and hence more likely to find representation in the right hemisphere. On this view, FS+ people have a more widespread representation of individual lexical items, and hence can access each word more readily and distinctly from syntactic processing than FS− people (Bever et al. 1987, 1989). This hypothesis would explain the relative ease of processing lexical items in FS+ people.

This interpretation is consistent with our recent finding that the age of the critical period differs as a function of familial handedness: FS+ deaf children show a younger critical onset age for mastery of ASL than FS− children (Ross and Bever 2004). This follows from the fact that FS+ people access the lexical structure of language more readily, and access syntactic organization less readily, than FS− people: FS+ children are acquiring their knowledge of language with greater emphasis on lexically coded structures, and hence depend more on the period during which vocabulary grows most rapidly (between five and ten years, itself possibly the result of changes in social exposure and emergence into early adolescence).

This lead to a prediction: lexical processing is more bilateral in FS+ right-handers than FS− right-handers, but syntactic processing is left-hemisphered for all right-handers. Recently, we tested this using fMRI brain imaging of subjects while they are reordering word sequences according to syntactic constraints or according to lexico-semantic relations between the words. We found suggestive evidence that the lexical tasks activated the language areas bilaterally in FS+ right-handers, but activated only the left-hemisphere areas in the FS− right-handers: all subjects showed strong left-hemisphere dominance in the syntactic tasks (Chan et al. in preparation). This confirms our prediction, and supports our explanation for Luria's original clinical observations. It also demonstrates that there is considerable lability in the neurological representation of important aspects of language. In an event-related potential (ERP) version of the Townsend et al. (2001) word probe study, we ~~have~~ found evidence for FS-mediated variability in the lateralization of the P200 ERP component, a possible marker for early lexical processing. FS+ ~~right-handers~~ people show a relatively large right-lateralization ~~of the component~~ response pattern compared to FS− ~~right-handers~~ people, supporting the hypothesis that lexical processing involves more bilateral function in FS+ (Hancock and Bever 2010).

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/13   Time:09:50:40
Filepath:d:/womat-filecopy/0001977625.3D403

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

There are reasons to believe that the genetics of handedness is multifaceted. Categorical phenotypes can be mapped to a continuous latent variable using a standard multifactorial threshold model (Falconer 1965). Accordingly, we have applied a general Bayesian multifactorial model to our collection of 3,000 family-handedness pedigrees to estimate the genetic effects of familial handedness in subjects (Hancock 2012). Emerging results from our laboratory using this measure, in conjunction with EEG measures, promise to identify familial handedness effects more robustly than previous behavioral studies.

For example, as mentioned above, ~~we have found a rightward shift~~ in P200 amplitudes associated with lexical access as a function of familial sinistrality in pure right-handed individuals (Hancock and Bever 2010). We have scaled the right-hemisphere P200 effect as a function of liability of each subject and found a significant correlation in right-handed subjects with the likelihood that they would have been left-handed (Hancock and Bever 2011): this lends initial support and validity to the liability measure and to its significance for functional processing of language. In addition, current results show a similar correlation for the standard EEG early left-negativity, found in response to sentences with a local ungrammaticality—the early negativity is more bilateral for FS+ and more left-lateralized for FS− subjects (Neville et al. 1991; Hahne and Friederici 2002; Sammler et al. in preparation).

Now consider the implications of these findings. The most immediate, which I won't belabor, is that it will enrich the clinical and theoretical study of neurolinguistics: almost no clinical or experimental investigations of neurolinguistic issues take the subjects' familial handedness into account, yet we now know that this can have profound impact on how language is neurologically organized.

The second implication involves the extent to which the "normal" neurological organization for language is not fixed. Prior to this, various cases of unique neurological organization for various components of language have called into question the idea that there is a single form of representation: these include a variety of reported cases of highly specific (and sometimes almost unbelievable) language deficits, cases of left-hemispherectomy in which the patients with a lone right hemisphere can grow up to be normal linguistically (Curtiss et al. 2001; Devlin et al. 2003) with normal developmental stages (Curtiss and Shaeffer 1997) as well as unique instances such as the in/famous hydrocephalic mathematician whose neocortex was a thin layer of tissue lining the skull (cf. Lewin 1980)—clearly the topology and connections of different cortical areas are very different in these cases from the norm. Even classic and recent studies call into question the unique location and function of a linguo-central structure such as Broca's and Wernicke's areas (Penfield

Comp. by: PG2846   Stage : Proof   ChapterID: 0001977625   Date:23/3/13   Time:09:50:40
Filepath:d:/womat-filecopy/0001977625.3D404

OUP UNCORRECTED PROOF – FIRST PROOF, 23/3/2013, SPi

and Roberts 1959; Bogen and Bogen 1976; Rogalsky and Hickok 2011). But people with familial left-handedness comprise roughly 40 percent of the population, so we cannot consign their unique behavioral and neurological structures to an odd minority.

A profound implication for language of all these considerations is the possibility that the existence of language is not causally dependent on any particular unique neurological organization. Rather, syntax especially is a computational type that recruits different neurological structures, originally evolved over ages for other modalities. On this view, the possibility for syntax emerges as a function of factors as yet undetermined: one possibility is the cognitive availability of propositional relations and categories, combined with an explosive growth in the number of lexical items that can externalize the internally represented concepts; such factors may interact with other principles of organization to result in the overt language structures (see, e.g., Nowak et al. 2002 for arguments of this sort). In this scheme, the syntactic architecture is represented neurologically via co-option and integration of different brain regions that are already adapted in other modalities for the type of computation that hierarchically structured language requires: they are felicitously connected to other areas that are also adapted for other types of language computations. Accordingly, there can be significant lability of how language will be represented in an individual's brain, if there is significant variability in how the computationally relevant areas function or are interconnected.

This approach will lead to a rich paradigm for the study of the relation between language and genetic factors. We can now use the familial handedness pedigree to predict the likelihood that a newborn will be left-handed: this gives us an important tool in tracking the simultaneous emergence of language in infancy along with the emergence of specific brain organization for language. That may clarify the extent to which language is shaped by universals of neurological maturation, and the extent to which its structure is independent of any particular neurological organization. The implication of that is consistent with the view that language as a biological system may be dispersed in the nervous system rather than dependent on specific locations (Chomsky 2000).

22.9 Conclusion

The more we study language with new tools of investigation, the more mysterious it becomes. I have suggested that the child's problem is vastly more difficult than "merely" figuring out how to combine words and morphemes:

the problem is how to isolate them in the first place from input that has already encoded, elided, and eliminated them. This heightens the salience of the idea that important properties of language come to us "for free," that is, as a function of language properties as "natural forms" (Hauser et al. 2006). ently, considerable interest has surfaced in these phenomena, exploring the extent to which language universals emerge as the result of natural laws applying to shape and connect the pieces that the child can recognize. This interesting line of thinking is further indirectly supported by our current work showing normal individual differences in the neurological organization for language.

old and new

Will language as an "organ," discussed in the initial quotation in this paper, turn out to be more like the immune system or skin than the liver?