

Linguistics: The Study of the Language Capacity and Its Functions

Diogenes

58(1–2) 20–34

Copyright © ICPHS 2012

Reprints and permission:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0392192112441910

dio.sagepub.com



Elizabeth Closs Traugott

Stanford University

Introduction

The capacity for language has for millennia been considered a defining property of humanity central to cultural development. In Judaeo-Christian mythology divine power is invested in language (*In the beginning was the Word*), and in the act of speaking (*Let there be light*). After Adam is created, beasts and fowl are created as his companions and God *brought them unto Adam to see what he would call them: whatsoever Adam called every living creature, that was the name thereof*. (Cited from *Holy Bible*, King James version, *John* 1.1, *Genesis* 1.3, and 2.19 respectively.) The act of naming empowered Adam, and differentiated him from the beasts of the earth. Likewise, in the Mayan sacred book, *Popul Vuh*, language is specifically linked with humans. When the first humans were fashioned out of corn *They talked and they made words. They looked and they listened*. (Text of the *Popul Vuh*, available at < <http://www.fiu.edu/~northupl/populvuh.html> > [accessed January 26th 2009]). More recently, in his presidential address to the Modern Language Association, Michael Holquist said “language—and the wisdom and community that are impossible without it—is a prerequisite, indeed it is arguably *the* prerequisite, for attaining the status of human being” (Holquist 2008: 570).

Probably the vast majority of people think very little about the components of language, the molecules, so to speak, of which it is made. They may be deeply concerned about language as a symbol of political identity, or about “correctness,” or fascinated by etymologies, but for the most part the capacity for language is taken for granted, just as we take our blood or our digestion for granted. The discipline of linguistics focuses on questions regarding the nature of the capacity for language (including sign language), the ways languages are structured, and the various functions of language in society. Above all, linguistics aims to find regularities and patterns, both within and across languages, in what may appear to be a heterogeneous and ultimately chaotic phenomenon. The perception of heterogeneity in language is reflected in the vast range of myths about the dispersion of languages, such as the story of the Tower of Babel. But in actuality there is much in common among the languages of the world.

Corresponding author:

Elizabeth Closs Traugott, Department of Linguistics, Stanford University, Stanford CA 94305-2150, USA

Email: traugott@stanford.edu

Linguistics is driven by testable hypotheses and conceptualized as a science. Languages and language in general are studied much as biology studies living species. Attention is paid to what is, not what “ought to be,” that is, it is descriptive rather than “prescriptive” or concerned with correctness. The history of the field is very old, and goes back at least to Panini, a Sanskrit grammarian of around the fourth century BCE, but in its current form the discipline took shape in the nineteenth century. Four “defining moments” when ideas in the air were synthesized may be noted in its development from the nineteenth century on. One was the schematization and modeling of systematic relationships among Indo-European languages in terms of a “family tree,” a concept akin to Darwin’s tree of life. (The concept of systematic relationships among languages originated in observations by Sir William Jones [1786] regarding similarities between Sanskrit and European languages like English and Greek. A number of ideas on the matter were systematized by Schleicher 1861–62. Schleicher acknowledged Darwin’s work, and likewise Darwin 1871 acknowledged Schleicher’s.)

The conclusion drawn from this work was the hypothesis that sounds change regularly, that is, without exception (except when borrowing interferes). Linguists began to view their field as a physical science. Another defining moment was the development of structuralist theories at the beginning of the twentieth century, for example, de Saussure’s contrast between *langue* ‘system’ and *parole* ‘speaking’ (“usage”), his privileging of the study of former, and his focus on the arbitrariness of the sign (for example, the word *cat* cannot be derived from anything inherent in felines or in our understanding of them). The third occurred in the mid-twentieth century, when Chomsky construed linguistics as the study of mind and brain. And the fourth “defining moment” to be identified here is the development in the later part of the twentieth century of extensive electronic data banks or “corpora,” ranging from ancient historical texts to contemporary recordings of conversation.

Linguistics has its roots in the study of grammar, philology, philosophy, anthropology, psychology, and pedagogy. From the mid-twentieth century on it has been closely associated with the cognitive sciences. With the advent of computers, linguistics has expanded into artificial intelligence. Medical applications have included speech and hearing pathology. More recently advances in neuroscience, neurobiology, and genetics have contributed to discussion of the evolution of language, and to what extent different aspects of language are located in the brain. Linguistics is therefore a multi-disciplinary field. (Aronoff and Rees-Miller, eds 2000 provide a useful survey of several aspects of the field at the end of the twentieth century.)

However, since the core question is whether language is unique to humans, it has firm roots in the humanities, and linguistics is institutionally defined as a humanities subject in many universities.

In what follows I will outline some of the questions and the answers to those questions that have been central to linguistic debate in the last half century, with particular attention to “the *raison d’être* of linguistics,” its present state and to emerging developments within it, as they pertain to the humanities. In a turn away from the behaviorism that characterized much of the earlier part of the twentieth century, and in publications starting in 1957 with *Syntactic Structures*, Chomsky focused attention on the hypothesis that the language faculty is genetically determined, in other words, innate, part of our DNA. He proposed three “basic questions” (Chomsky 1986: 3):

- (i) What constitutes knowledge of language?
- (ii) How is knowledge of language acquired?
- (iii) How is knowledge of language put to use?

In recent decades most work in linguistics has been devoted to answering these questions, explicitly or implicitly, either accepting the assumptions behind them, or challenging them. I start

by outlining the so-called formal generative approach, that accepts the assumptions, and seeks to answer the first two questions, then go on to so-called functional approaches that challenge the assumptions and seek especially to address the third question. I will pay particular attention to the question of how the link should be understood between covert meaning and overt form (spoken sounds, written graphs, or signs in sign languages). Views on this question are deeply embedded in views about the relative importance of different functions of language. Given the limits of space, I have had to be selective, and maximally opposing views have been drawn. In actual fact, there has recently been considerable convergence among many practitioners of formal generative and functional approaches, despite differences in assumptions.

Formal generative approaches

The assumptions behind questions (i)–(iii) are that “knowledge of language,” or “grammar” is something relatively fixed and definable, a “competence” acquired through the interaction of innate structure and input. (A highly readable introduction to the study of language from a largely generative approach is Pinker 1994. Penke and Rosenbach, eds 2007 discusses what counts as evidence for innateness in different theoretical approaches.) In the individual language-user competence pre-exists use, “communication and other special uses of language” (Chomsky 1986: 4). It is possible to give an explicit (“generative”) account of the “grammar,” which is a highly complex system. This system is considered to involve at a minimum abstract combinatory constraints on sentence structure (syntax), the sound system (phonology), and the meaning system (semantics), together with a set of units that give overt expression to the abstract structures (lexicon). “Grammar” in this sense is not a prescriptive enterprise, specifying what is “good” or “correct” language, but rather an account of what native speakers master without explicit instruction.

Construing communication and language use as derived from a system of knowledge is the direct result of seeking to identify what is uniquely human-specific about language. It is incontrovertible that communication is not human-specific. It is highly developed among other primates, as well as bees, and birds, to name only a few species. Likewise, construing meaning from context is not human-specific. In Chomsky and his colleagues’ view, what is human-specific is the ability to produce an infinite number of sentences with finite resources, to recognize covert differences between strings of words that appear to be similar, and to know that certain kinds of structural manipulation are not possible, all without appeal to context. Further, the function of human language is regarded as “a system for free expression of thought, essentially independent of ... instrumental purpose” (Chomsky 1980: 239). The capacity for language is not a product of culture, but of biology. On this view, communication is secondary to the language capacity.

Of fundamental importance in the formal generative approach to language, then, is our capacity to produce a sentence with no principled end other than the speaker’s ability to continue talking, as in:

(1) I knew that she knew that he knew how...

(1) illustrates “recursion,” the property that allows one sentence to be embedded in (or “subordinated to”) another. It follows that there is in theory no longest possible sentence, nor a largest set of possible sentences. Recursion is what enables us to “meta-represent,” in other words, to think and comment about thinking and commenting. These abilities have been called “creativity” in language. This creativity has little to do with what might be thought of as literary creativity, such as the ability to produce new metaphors, but rather is consistent with our mathematical ability to

make an infinite set of numbers based on the numerals 0–9 and a small number of operations like addition and multiplication.

The ability to recognize covert differences between strings of words that appear to be similar is illustrated by such examples as:

- (2) a. John expects Mary to like the dress.
b. John wants Mary to like the dress.

Native speakers of English know (without instruction) that (2a) can have a passive version as in (3a), but not (3b) (starred sentences are those considered by native speakers to be “ungrammatical” or not part of the system):

- (3) a. Mary is expected to like the dress.
b. *Mary is wanted to like the dress.

Likewise, the question forms in (4a, b) are possible, but not those in (4c, d):

- (4) a. Who is expected to like the dress?
b. What is Mary expected to like?
c. *Who is wanted to like the dress?
d. *What is Mary wanted to like?

This suggests that speakers have different “underlying representations” for *expect* and *want*, such that (2a) and (2b) roughly express (5a) and (5b) respectively:

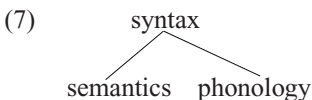
- (5) a. X expect Y [Y like Z]
b. X want [Y like Z]

It also shows that knowledge of how to construct a question depends not on the ability to extract any noun phrase from a string and question it, but rather to pay unconscious attention to abstract structures, and the different forms that embedding can take in a particular language.

In Chomsky’s view, it is abilities of this type that define our species-specific linguistic capacity. Of the three main structural properties of language: syntax, semantics, and phonology, syntax is therefore considered to be the core of the language capacity; meaning is derived from it, and sound is its expression. In *Syntactic Structures* Chomsky famously suggested that syntax and semantics are separate modules, as evidenced by the fact that:

- (6) Colorless green ideas sleep furiously

is syntactically entirely well-formed, but not semantically so. Not only are syntax, semantics, and phonology separate modules (and by hypothesis separately accessed in the brain), but they are autonomous from general cognition. So the relationship between covert meaning and overt expression is posited to be roughly as follows:



Although formal generative theories are by no mean homogeneous and have developed radically over time, one characteristic common to them is the assumption about what an explanatory theory entails. Explanations must be simple (the principle of Economy), and they must predict phenomena beyond those for which they were formulated. A third assumption until recently has been that explanation should be theory-internal, in other words, syntax should be explained in terms of syntax, semantics in terms of semantics, and phonology in terms of phonology. However, recently work on interfaces between all three components has shifted attention to the linguistic system as a whole. For example, to give a full account of the two possible interpretations of (8) that appear as (8a) and (8b):

- (8) It was the fox that stole the chicken
 a. The thing (we saw) was the fox that stole the chicken
 b. The fox (and not the dog) stole the chicken

it is necessary to show how the biclausal syntax interfaces with semantics (does *it* refer to the fox, as in (8a), or is it non-referring, as in (8b)?), and with phonology (only the contrastive reading (8b) has stress on *fox*).

In the sixties and seventies investigation of systematic patterns, and of differences between “surface strings” and “underlying structure” such as are illustrated by (2) and (8), and the concept of “creativity” as recursion, led to extensive work on stylistics. (e.g. Freeman, ed. 1970, 1981.) In 1973 Kiparsky posed questions such as

- (i) What patterns are relevant in poetry?
 (ii) What linguistic sames are relevant in poetry? (Kiparsky 1973: 234)

He argued that the language of poetry is essentially derivative of linguistic form, and challenged the hypothesis that a theory of poetic language or of meter is independent of a theory of the structure of the language used. For example, patterns of alliteration and rhyme depend on the sound patterns of the language. Old English alliteration (word initial rhyme) is associated with the word initial stress typical of Germanic languages. When stress came to be shifted toward the end of the word after the massive import of words from French, rhyme came to be preferred. Metrics should be viewed in terms of strong (s) and weak (w) positions in a line, rather than in terms of feet. For example, the iambic pentameter line consists of five pairs of w s positions; the stress unit that counts is one that in everyday speech can have primary stress. This is illustrated by a standard iambic line such as (9):

- (9) The cur/few tolls / the knell / of part/ing day (Gray)

Complex interactions with syntax and word-formation account for the fact that although disyllabic compounds like *first-born* tend to begin on the even syllable (10a), they can also occur in the odd position, as does *daylight* in (10b). However, lines like (10c) do not occur since *pallet* is not a compound, and does not have the right kind of stress pattern (Kiparsky 1975: 587).

- (10) a. When Ap/ril’s first-/born flowers, /and all / things rare (Shakespeare)
 b. Yond light / is not / daylight, / I know / it, I (Shakespeare)
 c. *As the / pallet / on which / we must / expire

Work of this type seeks, as in linguistics generally, not only to account for patterns that occur, but also to predict those that do not, and to find cross-linguistic, typological generalizations.

In the later part of the twentieth century the focus of generative syntax came to be increasingly on refining the abstract properties of innate structure, defined as UG (universal grammar, which is by hypothesis “hard-wired” in DNA). One hypothesis is that there is an invariant core (the “Principles” of the innate linguistic system) shared by all languages, e.g. all languages have nouns, verbs, phrases, and possibilities for embedding, and a set of choices or “Parameters.” The parameters account for differences among languages. So, for example, a Verb Phrase may be Object-Verb, or Verb-Object. The language learner acquires one or other setting of the parameter, depending on input. Each parameter is linked to a cluster of properties. In Object-Verb languages, the finite auxiliary verb is clause-final, while in Verb-Object languages it is clause-medial, as illustrated by the Old English example in (11a) and the Modern English equivalent in (11b) respectively.

- (11) a. aefter þaem þe hie þiss gesprečen haefdon (late 9th century, *Orosius*)
 after that they this said had
 b. after they had said this

An important current project is known as the “cartography of syntactic structures.” The objective is to delineate and characterize “the properties that, by virtue of holding of the (syntactic component of the) human language faculty, will be found to hold of every human language” (Kayne 2005: 3), and the constraints on the ways in which languages can differ, e.g. what are possible “determiners” (elements like *that, the, every*) in a language, which determiners does a language have, if it has them (Latin and contemporary Lithuanian do not have articles), and what correlates are there with presence or absence of those that are optional?

In conclusion, answers to the three basic questions in section 1 are characteristically as follows. Knowledge of language, as an abstract cross-linguistic capacity, is constituted by the essentially syntactic structure of the genetically endowed potential for language (UG). In this sense it is arbitrary. The child acquires knowledge of a particular language when exposed to language use; acquisition is the result of the interaction of UG with input. The main function of language is expression of thought, and representation of propositions.

Focus on the nature of UG has led to increasing alignment of generative grammar with cognitive science, artificial intelligence, and neuroscience, although deep connections persist with the humanities as comparative grammars of languages across the world have come to be written.

Functional approaches

From the beginning, formal generative approaches have been challenged by a different concept of language which privileges meaning, context, and diverse functions (see Van Valin Jr. 2001). Above all, the linguistic system is thought to be shaped and constrained by the functions it serves. It not only enables thought and representation of propositions and states of affairs (“knowing that”), but is also fundamentally instrumental (“knowing how”)—it enables us to reason, and to do things through use of questions, promises, directives, and so forth, all linguistic actions known as “speech acts.” (Work on speech acts originates in philosophy, see Austin 1962 and Searle 1969.) Speech acts take a certain form—usually the subject is the first person, and the tense is present. If someone says (12a) this is not just a statement about the future, but a commitment on their part to do something (pay taxes), and failure to fulfill the commitment can have significant consequences. But if they say (12b), in the past tense, this sentence does not function as a promise, only as a statement about one:

- (12) a. I promise to pay all my taxes.
 b. I promised to pay all my taxes.

Speech acts are performed by speakers and are ways by which they regulate social interaction. Therefore, for many researchers in this paradigm, the capacity for communication is as important as that for thought.

From this perspective, explanatory theories show how the various aspects of language (semantics, syntax, and phonology) interface with each other, and “external” factors such as reasoning, categorization, and perception are taken into account in addition to theory-internal factors. In particular, semantics and “pragmatic” (implied) meanings are invoked in work on syntax. No module is considered autonomous (and in some cases, no modules are postulated). Functional properties of language are conceived in terms of tendencies and continua, or of prototypes. Indeed, the word order parameters mentioned above were originally conceptualized as tendencies in use (Greenberg 1966), not as hard wired on–off choices in DNA.

Approaches of this type have come to be known as “functional,” and have been gaining considerable ground in recent years. Two strands of functional linguistics will be discussed here. One is associated with what is called “cognitive linguistics” and more recently “construction grammar.” Another is associated with structured variation in use, including discourse, in other words, in stretches of language longer than the sentence.

Cognitive linguistics

The basic hypothesis of cognitive linguistics is that the linguistic capacity is not autonomous and independent of cognition. Rather, it is symbolic. It is essentially a “pairing between a semantic structure and a phonological structure, such that one is able to evoke the other” (Langacker 2008: 5). Meaning is understood as conceptualization. This is true not only of words like *desk*, which refer to objects in the world, but also of grammatical notions such as noun, or subject.

Central to much work on cognitive linguistics is the observation that language is fundamentally metaphorical. Abstract concepts are mapped on to concrete expressions via stereotypical propositions such as ‘Language is a container’, ‘Argument is a building’, as evidenced by:

- (13) a. Try to *pack* more *thoughts into* fewer words. (Reddy 1993: 167)
 b. With the *groundwork* you’ve got, you can *construct* a pretty *strong* argument. (Lakoff and Johnson 1980: 98)

One of the more widely accepted hypotheses is that most linguistic expressions are based in the body and bodily experience. One particularly good example of “embodiment” is the way we conceptualize time, either as an entity coming toward us (*in the coming weeks*), or as something to which we move (*I am going to give a talk tomorrow*). In the history of languages, we find over and over again that terms for body parts come to be used as prepositions (e.g. terms for ‘belly’ come to be used to express the abstract location ‘in’), or that we conceptualize mental ability in terms of the body (e.g. one can be mentally ‘behind’, or ‘on top of’ things; see Heine, Claudi, and Hünemeyer 1991). Metaphors conceptually involve partial matching of perceived similarities; the thinking involved is analogical. In recent years the role of associative thinking or metonymy has also been emphasized. For example, cause and effect metonymies as in (14a) and controlled for controller metonymies in as in (14b) are pervasive in language (Kövecses 2006: Chapter 7):

- (14) a. She's my *pride and joy* (emotion for the cause of emotion)
 b. marry *money* (possessed for possessor)

The orientation to metaphor and metonymy has led to extensive discussion of the role of language in culture, and to proposals that language is both dependent on thought and also influenced by it. A contemporary example is the appearance of phrases like *the greening of the environment*, or *human resources* in the last few years, both expressions and concepts that played little role in the construction of ideologies until recently. Considering an older example, Wierzbicka hypothesizes that the concept of what is *reasonable* changed with the advent of the enlightenment. A 'reasonable' person in the sixteenth century was understood as one who was endowed with reason, but in the eighteenth century and beyond as one who is moderate and weighs evidence (cf. the philosopher Locke's insistence on empirical evidence). (Wierzbicka 2006: Chapter 4.)

Work of this kind serves as reminder that contemporary meanings cannot be projected onto texts from earlier times without considerable care, and that an ideologically neutral linguistic theory is impossible. (For discussion of methodologies for interpreting historical texts, see Fitzmaurice and Taavitsainen, eds 2007.)

From this perspective "the literary study of topoi (conventionalized kinds of [linguistic_{ECT}] content)" is construed as part of the study of cognition. "Cognitive poetics" seeks to understand for example, how Wallace Stevens's line *Death is the mother of beauty* is effective but *Death is the father/parent of beauty* would not be, because the latter is not grounded in idealized or prototypical cognitive models of kinship (Fabb 2001).

Again, as in the work on metrics discussed above, the basic insight is that literary language is not a separate domain from everyday language. Not being separate, it is subject to the same ethical questions, whether the issue is feminist readings, or conceptualization of criminality. For example, Fludernik (2005) discuss how spaces, bodies, and metaphors inform an understanding of how prison and criminality are conceptualized in literary texts at different periods.

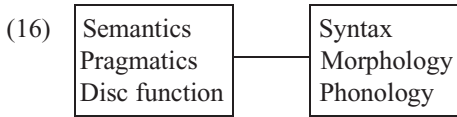
A particularly vibrant development out of cognitive linguistics has been that of "Construction grammar," synthesized by Goldberg. This brings together the hypothesis that language is essentially symbolic with the observation that many of the patterns found in it are "chunks" that appear to be learned piecemeal, not atomistically (Golberg 2006). The linguistic capacity is not restricted to producing and understanding the "core" structures that are the focus of generative work, but also embraces prefabricated routines and formulae ranging from idioms such as *by and large* to relatively marginal form-meaning pairings such as the incredulity construction illustrated by *Him a trapeze artist?!*, to highly ordinary ones like those in (15):

- (15) a. He sliced the bread.
 b. Pat sliced the carrots into the salad.
 c. Pat sliced Chris a piece of pie.
 d. Pat sliced the box open.

In (15) *slice* means 'cut with a sharp instrument', but the interpretation of the action depends on the "frame" in which it used: transitive in (15a) (the agent acts on the bread), caused-motion in (15b) (the agent caused the carrots to move), ditransitive in (15c) (the agent intended to cause someone to receive something), and resultative in (15d) (the agent caused a change of state). (Golberg 2006: 7.)

In construction grammar syntax is not especially privileged. Meanings are thought of in terms of properties of literal meaning (semantics), implied meaning (pragmatics), and discourse

function (e.g. information structure), and form is thought of in terms of properties of sentence structure (syntax), grammatical elements like the *–s* that signals subject-verb agreement as in *she talks* (morphology), and sound patterns (phonology). A model of the relationship between meaning and form (based on Croft 2001: 18) is given in (16) and may be contrasted with the one in (7). (“Disc function” is short for “Discourse function”).



Language is understood to be acquired throughout lifetime on the basis of input and general cognitive mechanisms. Little if any part of the language capacity is innate. Therefore knowledge of language does not exist *a priori*. It is built up by storage of patterns, some of which are item-specific, others of which are highly generalized and generalizable.

In conclusion, in construction grammar, and in cognitive grammar in general, language is viewed as a cognitive system. Propositional representation of thought is just one of the speech acts in which we engage on a daily basis. While some aspects of language are arbitrary, many are not, but are linked to bodily and other experiences. Creativity is accounted for by the free combination of constructions (or “chunks”).

The study of structured variation

Unlike in the research tradition starting with de Saussure, in which attention was paid almost exclusively to an (idealized) homogeneous language-user, another branch of functional linguistics is concerned with language as a heterogeneous capacity in which structured variation is essential to native-like command of a language. In many cases the nature of multi-dialectal communities and of language contact play center stage. In this work naturally occurring speech and writing is the dominant data, in contrast with the constructed, decontextualized examples based on introspection used in much of the research cited so far.

The correlates of variation were initially understood as social (associated with class membership) and stylistic (associated with more or less formal styles). Work on variation emphasized the ways in which variability arises out of change and is the basis of it (Weinreich, Labov, and Herzog 1968). Research on sociolinguistic aspects of language came to embrace narrative and interactive discourse. (For an overview, see Mesthrie et al. 2000.) Labov’s studies of the structure of narratives told by young African Americans (Labov 1972) became seminal for later work on narrative and its role in everyday conversation, identity building, medical interviews, legal trials, and a large variety of genres, including literary works. Likewise, Tannen’s work on talk among friends (Tannen 1984) set the agenda for much subsequent research on ways in which participants negotiate and thereby “co-construct” meaning in discourse. Tannen suggested that conversation is a system of pervasive sequential parallelism, akin to that noted by Jakobson (1966) in poetry. The excerpt in (17) from a conversation among four interlocutors about Schumann’s destroying his fingers for piano-playing by building a contraption to stretch them illustrates both co-construction of on-going talk and partial repetition (Tannen 1987: 591):

- (17) D: I read something in the newspaper.
 I won’t tell you.
 N: What contraption?

- S: I don't want to hear about it.
 D: You don't want to hear about it.
 L: Tell it, tell it.
 N: We want to hear about it.

(17) also illustrates a number of speech acts, and ways in which they can be used indirectly. D's *I won't tell you* appears to be a straightforward statement about intentions, but following immediately on the announcement of a news item, it implies that the newsworthy item is too gruesome to tell, so functions also as a warning about content. At the same time it serves as an invitation to be asked to tell the story after all. Modified repetition in the form of *You don't want to hear about it* reinforces this, and further serves to invite L's request that immediately follows: *Tell it, tell it*. The argument is that conversational, interactive spoken language is structured and has a "grammar" just as written language has; the systems are on a continuum, and not fundamentally different. Indeed, "spoken" vs. "written" is not a very useful dichotomy. After all, lectures are not read (even though etymologically *lect-* derives from Latin 'read'), and transcripts are written versions of spoken language. It is better to think of a continuum from "oral," largely spoken, styles like informal conversations to formal "literate" styles that are the outcome of training and literacy. E-mail to friends and blogs, though written, have often been said to tend toward the oral end of the continuum. However, since the online world has in many ways changed reasons for writing, the traditional categories may need rethinking, at least for that context (Baron 2008).

A fairly easy merger between the interests of philosophers of language in speech acts, and in pragmatics, and those of discourse analysts ensued (see Schiffrin, Tannen, and Hamilton, eds 2001), and contributed to extensive work on ways in which speakers and hearers build common ground, and coherent discourses (Clark 1996). This often involves invoking sets or scales. Two kinds of expressions that invoke such sets or scales are "information packaging" and "focus markers." Work on information packaging includes identification of the constraints on when it is possible to use the constructions illustrated in (18):

- (18) a. It was the sardines that Bill ate.
 b. Tennis I enjoy.

Neither utterance can easily occur conversation-initially. Both require or invoke a context in which there is a partially ordered set. In the case of (18a) the set is food that Bill ate, and the *It was* ("cleft") structure implies that sardines were the only food from the set that were eaten. In the case of (18b), the set is kinds of sport that the speaker likes. The requirement of a link to a prior discourse "anchor" that includes a set when preposing of a prepositional phrase occurs accounts for why *on the counter* in (19a) is coherent but *on the jacket* in (19b) is not unless a jacket in the kitchen has been a topic of discussion (Birnir and Ward 1998: 20):

- (19) a. I walked into the kitchen. On the counter was a large book.
 b. I walked into the kitchen. On the jacket was a large book.

The anchor (the kitchen) is stereotypically assumed to include such things as a counter. Jackets on the other hand are not part of our typical scenarios for kitchens. Work on focus markers seeks

to identify the meaning and function of little words like *only* and *even*. When speakers use them they evoke a scale and place an expression on it (low in the case of *only*, high in the case of *even*). Compare (20b, c) with (20a):

- (20) a. Jane will come to the party.
 b. Only Jane will come to the party.
 c. Even Jane will come to the party.

While (20a) implies no expectations, or evaluation regarding Jane vis-à-vis any other people who will come to the party, (20b) implies the speaker excludes all other members of the relevant set, and (20c) that it is unexpected that Jane would be a member of the set coming to the party. *Only* and especially *even* are subjective in the sense that they express the speaker's attitude. Both derive from non-subjective uses, and illustrate a widely-attested tendency for meanings of a particular expression to become more subjective over time and to function as markers of discourse-management. (See Traugott 2006.) Historically, focus marker *even* is related to *even(ly)* 'smoothly'; *only* originated in the sense 'uniquely'. If a word has multiple meanings, the more subjective one is usually the later one.

The access to vast quantities of naturally occurring data that was enabled by tape recordings and by electronic data bases led from the later nineteen-eighties on to an upsurge in attention to language use, both historically and in present-day English (UK, US, and other varieties). A useful list of some of the most important corpora available in the first years of this century can be found in McEnery and Gabrielatos 2006: 57–60. As in other areas of the humanities where digital data bases have become widely available, there is a concern that editions used may obscure original textual details. This is especially true for texts that pre-date printing, and above all for the Old English period, in which punctuation practices were very different. Modern editors have often imposed punctuation, thereby obscuring tricky questions about the complexity of particular sentences, identification of speakers, and so forth.

The corpora that have been compiled at various times since the 1950s have made it possible to do fine-grained statistical work on variation and change, including recent changes in the twentieth century. These have revealed that common perceptions, some of them supported by or induced by journalists and others who decry certain kinds of perceived changes, are in fact not borne out by empirical evidence. For example, it is often thought that the "simple" modals, like *must*, *may*, *can*, are declining and being replaced by "complex" semi-modals like *be going to*, *have to*. It is true that some fairly far-reaching changes have been occurring (and have been doing so ever since the modals first appeared around 1000 CE). In the twentieth century *shall*, *need*, and *ought to* have plummeted in use, while *may* and *might* have declined, but *will*, *can*, and *would* are still very robust. Use of *need to*, *want to*, and *be going to* increased in the twentieth century, but the surprising discovery is that all the complex semi-modals in aggregate occur less frequently than the single modal *will*. Therefore, it is clear that complex modals as a group are not replacing simple ones (Leech, Mair, Hundt, and Smith, 2009). Biber et al. (1999) is a major grammar of contemporary English based on a corpus of over 40 million words that seeks to account for differences between British and American English, both spoken and written, across conversation, news, fiction, and academic prose.

Frequency studies reveal patterns of usage in language that confirm the hypothesis of theories like construction grammar that much of language is routine, and formulaic, but is also used inter-actively and dynamically. Bybee and Hopper, for example, explore the hypotheses that:

“1. The distribution and frequency of the units of language are governed by the content of people’s interactions, which consist of a preponderance of subjective, evaluative statements, dominated by the use of pronouns, copulas [i.e. ‘*Be*’-verbs] and intransitive clauses.

2. The frequency with which certain items ... are used has a profound influence of the way language is broken up into chunks in memory storage, the way such chunks are related to other stored material and the ease with which they are accessed”. (Bybee and Hopper 2001: 3)

From this perspective, knowledge of language includes knowledge of frequency probabilities. Knowledge of language does not preexist use, but arises out of it, and out of experience with input. That is, language is shaped by use, whether production or parsing, and use continuously shapes or “structuralizes” language.

Approaches to evolution of language

The last research domain that will be touched on here is evolution. In the late nineteenth century the topic was banned from serious research since only speculation was possible. In recent decades, however, the topic has gained momentum. But it is still not possible to do much more than speculate. We know that *homo sapiens* has existed for a hundred and fifty to two hundred thousand years, and may have had language as we know it about fifty thousand years ago, both very short periods of time for biological evolution. As for records, they go back little over five thousand years. Two major hypotheses have been put forward, linked to views on the extent to which the language capacity may be innate, and what its prime function is. One is that the language capacity appeared abruptly, perhaps as a side-effect of some other ability; the other that it evolved piecemeal. (A range of views on evolution of language can be found in Wray, ed. 2002.)

Given his belief that the language capacity is species-specific, and serves the function of expressing thought, not of communication, it is not surprising that Chomsky has argued that the capacity for language does not antedate *homo sapiens*. He has associated it with the development of brain complexity and more recently, with the development of recursion (Hauser, Chomsky, and Fitch 2002). The hypothesis is that the language faculty “in the narrow sense” (i.e. that of UG), is too complex to have arisen as a partial system and developed adaptively. Counter-arguments focus on language as communication, and suggest that language evolved piecemeal and adaptively by biological selection. (See the response to Hauser, Chomsky, and Fitch 2002 in Pinker and Jackendoff 2005.) Yet others have argued that given the short time span involved in the hypothesized development of language, biological natural selection does not appear to be a viable hypothesis; rather, one may posit cultural adaptation involving interactive feedback, which results in language becoming easier to learn, and more structured over time. (Cf. Kirby, Smith, and Cornish 2008. Kirby 1999, however, advances a theory of natural selection in biological as well as historical timescales.) One plausible suggestion is that “the holistic processing strategy we use for achieving manipulative interaction may predate the emergence of single referential words and the analytic grammar that combines them” (Wray 2002).

Hopes for direct access to evidence for the genetic nature of language were raised when mutations of a gene called *FoxP2* were discovered in several generations of a family in the 1990s. Individuals with this mutation had significant speech difficulties. But they also had a large number of other motor difficulties, and the likelihood that this gene supports hypotheses about the uniqueness or even the innateness of the language capacity is very low (McAndrew n.d.). A more promising genetic link has been suggested with *ProtocadheriaXY*, a gene for cerebral asymmetry

(Crow 2002). Only homo sapiens has this asymmetry – the left hemisphere is dominant for non-holistic, atomistic language use.

The study of language evolution is still very much in its infancy. Although it has been pursued very largely in computational, mathematical, and neuroscientific research, it clearly has broad implications for humanities, as it seeks to provide evidence for whether the capacity for language is unique to humans, and how it developed in historical and cultural timescales.

Conclusion

I have been able to touch on only a small fraction of the kinds of issues that are at the center of linguistics. Whatever the researcher's perspective, the search is for pattern and regularity, and for ways to account for the relationship between micro-variations within language and larger macro-structures that underlie our capacity for language. Understanding how we acquire language, interpret, and produce it should give us insights into what language, the "prerequisite for attaining the status of human being," consists of, how it is used, how it has been used, and how it changes in both form and function.

Acknowledgement

Many thanks to Tine Breban for comments on an earlier draft.

References

- Aarts, B and McMahon, A, eds (2006) *The Handbook of English Linguistics*. Oxford: Blackwell.
- Aronoff, M and Rees-Miller, J, eds (2001) *The Handbook of Linguistics*. Oxford : Blackwell.
- Austin, J L (1962) *How to Do Things with Words*. Oxford: Oxford UP.
- Baron, N S (2008) *Always On: Language in an Online and Mobile World*. Oxford: Oxford UP.
- Biber, D, Johansson, S, Leech, G, Conrad, S and Finegan, E (1999) *Longman Grammar of Spoken and Written English*. Harlow, Essex: Pearson Education.
- Birner, B J and Ward, G (1998) *Information Status and Noncanonical Word Order in English*. Amsterdam: Benjamins.
- Bybee, J and Hopper, P (2001) "Introduction to Frequency and the Emergence of Linguistic Structure," in Bybee, J and Hopper, P (eds) *Frequency and the Emergence of Linguistic Structure*, pp. 1–24. Amsterdam: Benjamins.
- Chomsky, N (1957) *Syntactic Structures*. La Haye: Mouton.
- Chomsky, N (1980) *Rules and Representations*. New York: Columbia UP.
- Chomsky, N (1986) *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.
- Clark, H H (1996) *Using Language*. Cambridge: Cambridge UP.
- Croft, W (2001) *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford UP.
- Crow, T J (2002) "Protocadheriaxy: A Candidate Gene for Cerebral Asymmetry in Language," in Wray, A (ed.) *The Transition to Language*, pp. 93–112. Oxford: Oxford UP.
- Darwin, C (1871) *The Descent of Man, and Selection in Relation to Sex*. London: J. Murray.
- Fabb, N (2001) "Linguistics and Literature," in Aronoff, M and Rees-Miller, J (eds) *The Handbook of Linguistics*, pp. 446–465. Oxford: Blackwell.
- Fitzmaurice, S M and Taavitsainen, I, eds (2007) *Methods in Historical Pragmatics*. Berlin: Mouton de Gruyter.
- Fludernik, M (2005) "Metaphoric (Im)Prison(ment) and the Constitution of a Carceral Imaginary," *Anglia* 123(1): 1–25.

- Freeman, D C, ed. (1970) *Linguistics and Literary Style*. New York: Holt, Rinehart and Winston.
- Freeman, D C, ed. (1981) *Essays in Modern Stylistics*. London/New York: Methuen.
- Goldberg, A E (2006) *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford UP.
- Greenberg, J H (1966) "Some Universals of Language with Particular Reference to the Order of Meaningful Elements," in Greenberg, J H (ed.) *Language Universals, with Special Reference to Feature Hierarchies*, pp. 73–113. The Hague: Mouton².
- Hauser, M D, Chomsky, N and Fitch, W T (2002) "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" *Science* 292(5598): 1569–1579.
- Heine, B, Claudi, U and Hünemeyer, F (1991) *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- Holquist, M (2008) "Presidential Address 2007: The Scandal of Literacy," *Proceedings of the Modern Language Association* 123(3): 568–579.
- Jakobson, R (1966) "Closing Statement: Linguistics and Poetics," in Sebeok, T A (ed.) *Style in Language*, pp. 35–377. Cambridge, MA: MIT Press.
- Jones, W (1786) "Third Anniversary Discourse: On the Hindus," *Asiatick Researches* 1: 415–431.
- Kayne, R S (2005) "Some Notes on Comparative Syntax, with Special Reference to English and French," in Cinque, G and Kayne, R S (eds) *The Oxford Handbook of Comparative Syntax*, pp. 3–69. New York: Oxford UP.
- Kiparsky, P (1973) "The Role of Linguistics in a Theory of Poetry," *Daedalus* 102(3): 231–247.
- Kiparsky, P (1975) "Stress, Syntax, and Meter," *Language* 51(3): 576–616.
- Kirby, S (1999) *Function, Selection, and Innateness: The Emergence of Language Universals*. Oxford: Oxford UP.
- Kirby, S, Smith, K and Cornish, H (2008) "Language, Learning and Cultural Evolution: How Linguistic Transmission Leads to Cumulative Adaptation," in Cooper, R. and Kempson, R. (eds) *Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution*, pp. 81–119. London: University of London.
- Kövecses, Z (2006) *Language, Mind, and Culture: A Practical Introduction*. Oxford: Oxford UP.
- Labov, W (1972) "The Transformation of Experience in Narrative Syntax," in Labov's *Language in the Inner City*, pp. 354–396. Philadelphia: University of Pennsylvania Press.
- Lakoff, G and Johnson, M (1980) *Metaphors We Live By*. Chicago: University of Chicago Press.
- Langacker, R W (2008) *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford UP.
- Leech G, Hundt M, Mair C and Smith N (2009) *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.
- McAndrew, A (n.d.) "FOXP2 and the Evolution of Language," *Molecular Biology*, www.evolutionpages.com/FOXP2_language.htm.
- McEnery, T and Gabrielatos, C (2006) "English Corpus Linguistics," in Aarts, B and McMahon, A (eds) *The Handbook of English Linguistics*, pp. 33–71. Oxford: Blackwell.
- Mesthrie, R, Swann, J, Deumert, A and Leap, W L, eds (2000) *Introducing Sociolinguistics*. Philadelphia: Benjamins.
- Penke, M and Rosenbach, A, eds (2007) *What Counts as Evidence in Linguistics: The Case of Innateness*. Amsterdam: Benjamins.
- Pinker, S (1994) *The Language Instinct: How the Mind Creates Language*. New York: William Morrow.
- Pinker, S and Jackendoff, R (2005) "The Faculty of Language: What's Special About It?" *Cognition* 95(2): 201–236.
- Reddy, M J (1993) "The Conduit Metaphor: A Case of Frame Conflict in Our Language about Language," in Ortony, A (ed.) *Metaphor and Thought*, pp. 284–384. Cambridge: Cambridge UP².
- Saussure, F de (1971) *Course in General Linguistics*, trans. Roy Harris. Chicago: Open Court.
- Schiffrin, D, Tannen, D and Hamilton, H E, eds (2001) *The Handbook of Discourse Analysis*. Oxford: Blackwell.

- Schleicher, A (1861–62) *Compendium der vergleichenden Grammatik der Indogermanischen Sprachen*. Weimar: Hermann Böhlau.
- Searle, J R (1969) *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge UP.
- Tannen, D (1984) *Conversational Style: Analyzing Talk among Friends*. Oxford: Oxford UP.
- Tannen, D (1987) “Repetition and Variation as Spontaneous Formulaicity in Conversation,” *Language* 63(3): 574–605.
- Traugott, E C (2006) “The Semantic Development of Scalar Focus Modifiers,” in van Kemenade, A and Los, B (eds) *The Handbook of the History of English*, pp. 335–359. Oxford: Blackwell.
- Turner, M (1987) *Death is the Mother of Beauty: Mind, Metaphor, Criticism*. Chicago: University of Chicago Press.
- Van Valin, R D Jr (2001) “Functional Linguistics,” in Aronoff, M and Rees-Miller, J (eds) *The Handbook of Linguistics*, pp. 319–336. Oxford: Blackwell.
- Weinreich, U, Labov, W and Herzog, M (1968) “Empirical Foundations for a Theory of Language Change,” in Lehmann, W P and Malkiel, Y (eds) *Directions for Historical Linguistics: A Symposium*, pp. 97–195. Austin: University of Texas Press.
- Wierzbicka, A (2006) *English: Meaning and Culture*. Oxford: Oxford UP.
- Wray, A (2002) “Dual Processing in Proto-Language: Performance without Competence,” in Id. (ed.) *The Transition to Language*, pp. 113–137. Oxford: Oxford UP.