1

LANGUAGE STRUCTURE IN ITS HUMAN CONTEXT: NEW DIRECTIONS FOR THE LANGUAGE SCIENCES IN THE TWENTY-FIRST CENTURY

William Croft

The science of language in the twenty-first century is likely to expand its scope compared to that of the twentieth century. The twentieth-century science of language focused its attention on the analysis of language structure: the sound system of a language (**PHONOLOGY**) and its grammatical system (**MORPHOLOGY** and **SYNTAX**). The analysis of linguistic structure, or form, is central to the science of language. After the middle of the twentieth century, however, greater attention was placed on the relationship between language form and its psychological and social context.

The analysis of linguistic structure will remain central to the science of language. However, understanding language in context will undoubtedly be a crucial feature of language science in the twenty-first century. This essay focuses on the basic principles that have emerged in research on language in its social and cognitive context, the ways that this context constrains language structure and use, and the new directions in research implied by the integration of language structure and context. This essay is necessarily selective in the topics covered, and the selection represents a particular way to integrate language form and its context. It also brings together theories that have originated in philosophy, psychology, and sociology, as well as different branches of linguistics. Such effort is necessary in order to treat language as a unitary phenomenon, and also to relate central questions of linguistic analysis to other scientific domains. Language structure cannot be fully understood without situating it with respect to current theories of joint action, social cognition, conceptualization of experience, memory and learning, cultural transmission and evolution, shared knowledge and practice in communities, and demographic processes in human history.

WHY TALK? THE PRAGMATICS OF LANGUAGE

Why do we talk? Why does language exist? It is only by answering these questions that we can understand how language fits in its context. The answer is that language plays an essential role in social interaction, fundamentally at the level of joint action between two or more individuals (Clark 1996; Tomasello, 2008). What makes a joint action joint is that it is more than just the sum of individual actions performed by separate persons; in particular, each individual involved must take into consideration the other individual's beliefs, intentions, and actions in a way that can be described as cooperative. A shared cooperative activity between two individuals can be defined in terms of a set of attitudes held by the cooperating individuals and as a way of carrying out the individual action (Bratman 1992). The attitudes are as follows:

(a) Each individual participant intends to perform the joint action. That is, each participant's intention is not directed simply toward his/her individual action but toward the joint action that is carried out by both participants together.

(b) Each participant intends to perform the joint action in accordance with and because of each one's meshing subplans. That is, each participant's individual actions are intended to mesh with the other participant's actions in order to successfully achieve the joint action.

(c) Neither participant is coercing the other.

(d) Each participant has a commitment to mutual support. That is, each one will help the other to carry out the subplans; each participant is thus responsible for more than just execution of his/her own subplan.

(e) All of (a)–(d) are common ground, or shared knowledge between the individuals. The concept of common ground plays a central role in understanding the function of language in social interaction; it is discussed more fully toward the end of this essay.

Finally, in addition to these mental attitudes on the part of the participants, there must be mutual responsiveness in action. That is, the participants will coordinate their individual actions as they are executed in order to ensure that they mesh with each other and, hence, that the joint action will be successfully carried out (to the best of their abilities). Coordination is essential in carrying out joint actions successfully, and this is where language plays a central role in joint actions.

The social cognitive abilities necessary for shared cooperative activity appear to be unique to humans, providing what Michael Tomasello (2008) calls the social cognitive infrastructure necessary for the evolution of the capacity for modern human language. Other species than humans have a capacity for imitative learning of complex vocalizations (see ANIMAL COMMUNICATION AND HUMAN LANGUAGE). This has not been sufficient to lead to the evolution of human-like language among these species. Nonhuman primates have the ability to plan actions and to recognize regularities in behavior of other creatures, enough to manipulate their behavior. These abilities are preconditions for executing complex actions such as joint actions, but they are not sufficient for doing so.

Research on primate behavior in natural and experimental settings suggest that some primates even have the ability to recognize conspecifics as beings with intentional states like themselves in some circumstances (Tomasello, 2008; this ability develops in humans only at around nine months of age). Nevertheless, it has not been demonstrated that nonhuman primates have the ability to engage in shared cooperative activity

as already defined. Tomasello (ibid.) suggests that in particular, helpfulness, Michael Bratman's condition (d), may be critical to the evolution of the ability to carry out joint actions.

The final condition for joint action is that the individual actions must be coordinated in accordance with the shared attitudes of the participants. Any joint action poses coordination problems between the participants (Lewis 1969). Any means that is used to solve the coordination problem on a particular occasion is a coordination device. There are various coordination devices that human beings use to solve the coordination problems of joint actions, of which the simplest is joint attention to jointly salient properties of the environment (Tomasello 1999, 2003). But by far the most effective coordination device is for the participants to communicate with each other: By communicating their mental states, the participants greatly facilitate the execution of any joint action.

COMMUNICATION is itself a joint action, however. The speaker and hearer must converge on a recognition of the speaker's intention by the hearer (see COMMUNICATIVE INTENTION; see also COOPERATIVE PRINCIPLE). This is H. Paul Grice's definition of meaning ([1948] 1989), or Herbert Clark's informative act (Clark 1992; see the next section). And this joint action poses coordination problems of its own. The essential problem for the joint action of communication is that the participants cannot read each other's minds. Language is the primary coordination device used to solve the coordination problem of communication, which is in turn used to solve the coordination problem for joint actions in general. Indeed, that is the ultimate purpose of language: to solve the coordination problem for joint actions, ranging from the mundane to the monumental (Clark 1999). This fact is essential for understanding the structure of discourse and the linguistic expressions used in it, as Clark (1992, 1996) has shown for many aspects of conversational interaction, and it also accounts for many fundamental properties of linguistic structure.

LANGUAGE, COMMUNICATION, AND CONVENTION

A language can be provisionally described as a conventional system for communication (this definition is modified later in this section). David Lewis (1969) and Clark (1996, Chapter 5) define convention as follows:

(i) A regularity in behavior

(ii) that is partly arbitrary (that is, we could have equally chosen an alternative regularity of behavior),

(iii) that is common ground in the community,

(iv) as a coordination device

(v) for a recurrent coordination problem.

In other words, conventions can emerge when members of the community have shared knowledge that a certain repeated behavior can act among them as a coordination device for a recurrent coordination problem. This definition of convention is general: It applies to conventions such as shaking hands (or kissing on the cheek) for greeting, or driving on the right (left) side of the road. The definition also applies straightforwardly to language: A string of sounds (i.e., a word or morpheme, such as *butterfly*) or a grammatical construction (such as the Modifier-Head construction for English noun phrases) emerges as a convention when it becomes a regularly used means for solving the recurrent coordination problem of referring to a specific experience that is to be communicated.

Linguistic convention actually operates at two levels: the grammatical level of words and constructions, at which the speaker's intentions are formulated; and the phonological level of the articulation and perception of the sounds that make up the grammatical units. This is the phenomenon described as duality of patterning in language (Hockett 1960). One could imagine in principle that linguistic convention possessed only one level: perceivable sounds (or **GESTURES** or written images, depending on the medium), corresponding to part (i) in the definition of convention, that directly conveyed the speaker's intention (the recurrent coordination problem) as a whole, corresponding to part (v) in the definition of convention. These exist in interjections with specific functions such as *Hello* and *Thanks*. However, most linguistic expressions are complex, consisting of discrete, meaningful units. Complex linguistic expressions evolved for two reasons: First, the number of different speaker intentions to be communicated grew to be indefinitely large; and second, a speaker's intended message came to be broken down into recurrent conceptual parts that could be recombined to produce the indefinite variety of messages.

Again, one could imagine that each conventional linguistic unit consisted of a unique sound (gesture, image). But languages have distinct meaningful units that are made up of different sequences of the same sounds: *bat*, *sat*, *Sam*, *same*, *tame*, *time*, etc. This system has evolved for the same two reasons: the increasing number of meaningful units (even the recurring ones) necessary to convey the indefinitely large number of speaker intentions, and an ability to break down a sound signal (or gesture, or image) into parts that can be recombined as a sequence of sounds (or gestures or images). Thus, the duality of patterning characteristic of human language has evolved to accommodate the huge number of speaker intentions that people want to convey, and to exploit the facts that intentions can be broken down into recombinable conceptual units and that the medium of expression can be broken down into recombinable units as well.

Language is therefore a joint action that operates simultaneously at four levels (Clark 1996). The higher-numbered levels are dependent on the lower-numbered levels; the individual actions of the interlocutors are given in italics:

(4) *proposing* and *taking up* a joint project (joint action);

(3) *signaling* and *recognizing* the communicative intention;

(2) *formulating* and *identifying* the proposition;

(1) *producing* and *attending* to the utterance.

The highest level corresponds to the **ILLOCUTIONARY FORCE** in speech-act theory (Austin 1962); the next level to Gricean meaning, or the informative act (Clark 1992); the next level to the propositional act (Searle 1969); and the lowest level to the utterance act (Austin 1962; Searle 1969). Each level enables the level(s) above it, and succeeds only if the level(s) below has been successfully achieved (e.g., one cannot recognize the

communicative intention if one did not pay attention to the utterance produced).

THE INCOMPLETENESS OF CONVENTION

The model of language as joint action describes the social cognitive system that must have evolved in the human species for modern human language to have emerged. It describes what appears to be a stable system that led to the emergence of highly complex cooperative activity among humans, namely, what is called society or culture. But it is not a complete picture of the nature of language in social interaction.

Linguistic convention can function as a coordination device for communication because there are recurrent coordination problems in communication: People have repeatedly wished to convey similar intentions formulated in similar concepts. Convention, linguistic or otherwise, is a regularity of behavior that emerges in a community or society. But convention must emerge from previous successful communication events where a convention did not previously exist. In other words, there must be a precedent: You and I use a coordination device because we used it before (or observed it used before), and it worked. Following a precedent is a coordination device, but it is not (yet) convention; it is based not on regular behavior that is mutually known in the community but only on previous successful uses that we are aware of (Lewis 1969).

Following a precedent cannot be the ultimate root of convention either. It always requires a successfully coordinated communicative act as a precedent. The ultimate coordination device is joint salience: Each participant can assume that in a particular situation, certain features are salient to both participants (Lewis 1969). Joint salience is possible because humans have the social cognitive capacity for joint attention to their environment (Tomasello 2003). Joint attention forms a basis for common ground, as discussed later in this article.

Linguistic convention, however, is not perfect; it does not trump or replace the nonconventional coordination devices of precedent and joint salience in the act of communication. This is partly because of the kind of conventions found in language, and partly because of the nature of convention itself.

Linguistic conventions are incomplete because of the phenomena of indexicality and **AMBIGUITY** (Clark 1996). A linguistic convention such as *hat* or *find* represents a type, but on a particular occasion of use, we often intend to convey a particular token of the category. Thus, *I found the hat* communicates a particular taking event involving a specific hat. In order to identify which finding event and which hat, the interlocutors must rely on joint salience in the context, facilitated in part by the past tense of *find* and the article *the* combined with *hat*, to coordinate successfully on the right finding event and the right hat. Linguistic shifters, such as the pronoun *I*, more explicitly require joint salience, namely, who is the speaker in the context. Proper names denote tokens, but even a proper name such as *William Croft* may be (and is) used for more than one individual, for example, the contemporary linguist and the English Baroque musical composer.

Most words are also highly ambiguous; that is, the same regularity of behavior is used as a coordination device to solve different recurrent coordination problems. For example, *patient* is ambiguous between the linguistic semantic role (*The patient in sentence 25 is Roland*) and a role in the domain of medicine (*The patient in room 25 is Roland*). Linguistic convention alone cannot tell which meaning is intended by the speaker. Only joint salience, provided in the example sentences by the meanings of the other words and the broader context of conversation, will successfully solve the coordination problem of what is meant by *patient*.

Indexicality and ambiguity are so pervasive in language that no utterance can be successfully conveyed without recourse to nonconventional coordination devices. But convention itself is also incomplete. This is because every situation being communicated is unique and can be construed as the recurrence of different coordination problems. The simplest example of this phenomenon is that different words can be used to describe the current situation, each representing a different construal of the current situation in comparison to prior situations. For example, one can refer to an individual as *the prime minister, Tony Blair, the Labour Party leader, my friend, that guy, he*, etc.; each expression construes reference to the current person as the recurrence of a different coordination problem.

The need to use nonconventional coordination devices as well as linguistic convention in communication is not generally a problem for successful joint actions by cooperative human beings. However, in some contexts, successful coordination is quite difficult. For example, scholarly discourse on abstract theoretical concepts often leads to alternative construals of what is intended by particular scholars. What do we take Plato to have meant? This changes over time and across persons. Alternative construals, not always accurately described as "misunderstandings," occur in more everyday circumstances as well, as readers can verify for themselves.

In addition, human beings are not always cooperative. The complexity of language as joint action here leaves open many possible means of language abuse. For example, lying abuses linguistic convention in its role of helping coordinate a shared cooperative activity, namely, coming to a shared belief. Other types of language abuse exploit nonconventional coordination devices. For example, in one lawsuit, the courts ordered a government agency to destroy certain *documents*, intending the term to denote their information content; the agency destroyed the *documents*, that is, the physical objects, after making copies of them (Bolinger 1980). Here, the ambiguity of *documents* requires recourse to joint salience, but the agency abused this nonconventional coordination device (the lawsuit was about privacy of information). Finally, the fact that a current situation can be construed as an instance of different recurrent coordination problems leads to alternative framings of the situation, such as referring to an entity as a *fetus* or an *unborn baby*. These alternative framings bias the conceptualization of the current situation in ways that invite certain inferences and courses of action, rather than others.

THE LINGUISTIC SYSTEM IN CONTEXT

In the preceding sections, language is described as a conventional system for communication, and the role of convention

in language and of language in communication was discussed. In this section, the linguistic system is described in broad outline. Linguistic structure has been intensively studied over the past century ever since Ferdinand de Saussure inaugurated the modern analysis of linguistic structure, *Structuralism* (Saussure [1916] 1966). This section focuses on those aspects of linguistic structure that are generally agreed upon and shows the extent to which they emerge from the principles that have been presented in the preceding section.

The most fundamental structuralist principle is the centrality of the linguistic sign or symbol, that is, the notion that language pairs form and meaning, and that particular linguistic forms convey particular meanings. This principle fits directly with the definition of convention. The regularity in behavior in part (i) of the definition of convention is the expression of a linguistic form by a speaker; the recurrent coordination problem in part (v) of the definition is the communication of a meaning between the interlocutors.

Also central to the structural analysis of language is the arbitrariness of the linguistic sign. That is, arbitrariness exists in the particular form and meaning that are paired. This conforms with part (ii) of the definition of convention, namely, that the convention is partly arbitrary. Arbitrariness is usually defined in structuralist analysis as the principle that one cannot entirely predict the form used from the meaning that is intended. From a communicative point of view, arbitrariness means that another choice could have served approximately equally well. For example, the choice of producing the string of sounds *butterfly* for a particular meaning could have been replaced with the choice of producing the string of sounds *Schmetterling* – a choice made by members of the German speech community. Two different choices are communicatively equivalent in that neither choice is preferred for the meaning intended – and that is usually because the choice of one expression over the other is arbitrary in the structuralist sense.

Another principle that can be traced back to Saussure is the distinction between the paradigmatic and syntagmatic contrast of linguistic units. In a complex (multiword or multimorpheme) grammatical construction, such as *The cat sat on the mat,* each word enters into two different types of contrast. For example, the first word *the* contrasts with the word *cat* in that *the*'s role in the construction (determiner) contrasts with *cat*'s role (head noun). This is a syntagmatic contrast. But *the* also contrasts with another possible filler of the same role in the construction, such as *a* in *A cat sat on the mat*; and *cat* contrasts with *hamster, parakeet*, etc. in the same way. These are paradigmatic contrasts.

More recent grammatical theories represent paradigmatic contrast in terms of a set of elements belonging to a grammatical category. Thus, *the* and *a* belong to the category *determiner,* and *cat, hamster, parakeet,* etc. belong to the category *noun*. Syntagmatic contrasts are represented by contrasting roles in the syntactic structure or constructions used in the utterance. For example, the determiner category is functioning as a modifier of the noun category in a noun phrase construction. Finally, the syntagmatic–paradigmatic distinction also applies to phonology (sound structure): Paradigmatic contrast is represented by phonological categories, and syntagmatic contrasts by the phonological structure of words and larger prosodic units.

The syntagmatic–paradigmatic distinction is the most basic way to describe the fact that the linguistic system allows a (re-) combination of meaningful units in different ways. The adaptive motivation for the emergence of such a communication system was described previously: The number of intentions to be communicated is so great that a set of simple (atomic) symbols will not suffice, but experience is such that it can be broken down into recurrent parts for which conventional linguistic expressions can develop. The same motivations gave rise to the syntagmatic–paradigmatic distinction in phonology as well.

Paradigmatic principles of structure in grammar and phonology are represented in terms of linguistic categories, phonological and grammatical. These abstract linguistic categories can be mapped onto the substantive categories of the actual phonetic realization (for phonology) and of utterance meaning (for grammar). Linguistic typology (Comrie 1989; Croft 2003), which takes a cross-linguistic perspective on grammatical analysis, has demonstrated that the ways in which phonological categories are mapped onto phonetic space, and grammatical or lexical categories are mapped onto conceptual space, are not unlimited. For example, phonetic similarities and conceptual similarities constrain the structure of phonological and grammatical categories, respectively.

Syntagmatic principles of structure are represented in various ways, but all such representations reflect another basic principle, the hierarchical organization of the structure of utterances. Sentences are organized in a hierarchical structure, representing groupings of words at different levels. So *The cat sat on the mat* is not just a string of roles that contrast syntagmatically, as in [Determiner Noun Copula Preposition Determiner Noun]. Instead, it is a set of nested groupings of words: [[Determiner Noun] [Copula] [Preposition [Determiner Noun]]]. The nested groupings are frequently represented in a variety of ways, such as the syntactic trees of phrase (constituent) structure analysis. They can also be represented as dependency diagrams (for example, the determiner is related to the noun as its modifier, which in turn is related to the copula as its subject), and representations combining constituency and dependency also exist.

The structure of a construction often appears to be motivated, though not entirely predicted, by the structure of the meaning that it is intended to convey. For example, the syntactic groupings in [[*The cat*] *is* [*on* [*the mat*]]] are motivated semantically; *the* in *the cat* modifies *cat* semantically as well as syntactically (indicating that the cat's identity is known to both speaker and hearer). The (partial) motivation of syntactic structure by its meaning is captured by general principles in different theories. These principles can be described as variants of the broader principle of diagrammatic iconicity (Peirce 1932): roughly, that the abstract structure of the linguistic expression parallels the abstract structure of the meaning intended, to a great extent. It is difficult to evaluate the structure of meaning independently of the structure of linguistic form. However, different speech communities settle on a similar range of constructions to express the same complex meaning – the regularities discovered in linguistic **TYPOLOGY** (see, for example, the studies

published in Typological Studies in Language and the Oxford Studies in Typology and Linguistic Theory). This fact suggests that there are regularities in the meaning to be conveyed that are then reflected in the grammatical constructions used to express them.

GRAMMAR AND THE VERBALIZATION OF EXPERIENCE

The preceding sections have described the general context of language use and the basic principles of language structure. The grammars of particular languages conform to the basic principles of language structure in the preceding section. But the grammars of particular languages, while diverse in many ways, are similar to a much greater degree than would be predicted from the general principles in the preceding section, or even the context of language use described in the earlier sections. For example, all languages have structures like clauses in which some concept (prototypically an action concept, usually labeled a verb) is predicated on one or more concepts that are referred to (prototypically an object or person, usually labeled a noun). The noun-like expressions are in turn organized into phrases with modifiers. Clauses are related to each other by varying degrees of grammatical integration. Certain semantic categories are repeatedly expressed across languages as grammatical inflections or "function words" (e.g., articles, prepositions, auxiliaries) that combine with the major lexical categories of words in sentences.

These universal patterns in grammar are attributable to the way that experience is verbalized by human beings. The fundamental problem of verbalization is that each experience that a speaker wishes to verbalize is a unique whole. But a linguistic utterance is unlike an experience: An utterance is broken down into parts, and these parts are not unique; they have been used before in other utterances. (This latter point is the fact of convention; a particular linguistic form is used regularly and repeatedly for a recurrent coordination problem.)

The process by which the unique whole of experience is turned into a linguistic utterance made up of reusable parts has been described by Wallace Chafe (1977). The first step is that the speaker subchunks the experience into smaller parts, each also a unique Gestalt similar in this way to the original experience. The subchunking process may be iterated (in later work, Chafe emphasizes how consciousness shifts from one chunk to another in the experience to be verbalized). A subchunk of the experience is then propositionalized; this is the second step. Propositionalizing involves breaking up an experience by extracting certain entities that are (at least prototypically) persistent, existing across subchunks. These entities are the referents that function as arguments of the predicate; the predicate is what is left of the subchunk after the arguments have been separated. Propositionalizing therefore breaks down the experience into parts – arguments and the predicate – that are not of the same type as the original experience (i.e., not a Gestalt).

Once the whole has been broken down into these parts, the parts must be categorized, that is, assigned a category that relates the parts of the current experience to similar parts of prior experiences. Categorizing is the third step in the verbalization process. These categories are what are expressed by *content words*, such as nouns and verbs. In this way, the speaker has transformed the unique whole of the original experience into parts that can be expressed by language.

This is not the end of the verbalization process. Content words denote only general categories of parts of the experience to be verbalized. In order to communicate the original experience, the speaker must tie down the categories to the unique instances of objects, events, and so forth in the experience, and the speaker must assemble the parts into a structure representing the original whole that the speaker intends to verbalize. That is to say, corresponding to the categorizing step in verbalizing the parts of the experience, there is a particularizing step that indicates the unique parts; and corresponding to the steps of propositionalizing and subchunking are integrative steps of structuring and cohering, respectively (Croft 2007). These latter three steps give rise to grammar in the sense of grammatical constructions, inflections, and particles, and the semantic commonalities among grammatical categories across languages.

The particularizing step takes a category (a type) and selects an instance (token) or set of tokens, and also identifies it by situating it in space and time. For object concepts, selecting can be accomplished via the inflectional category of number, and via the grammatical categories of number and quantification (*three books, an ounce of gold*). For action concepts, selecting is done via grammatical aspect, which helps to individuate events in time (*ate* vs. *was eating*), and via agreement with subject and/or object, since events are also individuated by the participants in them (*I read the paper* and *She read the magazine* describe different reading events). Objects and events can be situated in space via deictic expressions and other sorts of locative expressions (*this book, the book on the table*). Events and some types of objects can be situated in time via tense and temporal expressions (*I ate two hours ago; ex-mayor*). Events and objects can also be situated relative to the mental states of the interlocutors: The article in *the book* indicates that the particular object is known to both speaker and hearer, and the modal auxiliary in *She should come* indicates that the event exists not in the real world but in the attitude of obligation in the mind of the speaker.

The structuring step takes participants and the predicated event in a clause and puts them together, reassembling the predicate and the argument(s) into the subchunk from which they were derived by propositionalizing. Grammatically this is a complex area. It includes the expression of grammatical relations in what is called the argument structure of a predicate, so that *She put the clock on the mantle* indicates which referent is the agent (the subject), which the thing moved (the object), and which the destination of the motion (the prepositional phrase). But it also includes alternative formulations of the same event, such as *The clock was put on the mantle* (the passive voice construction) and *It was the mantle where she put the clock* (a cleft construction). The alternative constructions function to present the information in the proposition in different ways to the hearer, depending on the way the discourse is unfolding; they are referred to as information structure or discourse function.

Finally, the cohering step takes the clauses (subchunks) and reassembles them into a whole that evokes the original whole

experience for the hearer. This step can be accomplished by various clause-linking devices, including subordination of various kinds, coordination, and other clause-linking constructions found in the world's languages. Coherence of clauses in discourse is also brought about by discourse particles and **REFERENCE TRACKING**, that is, grammatical devices, such as pronouns or **ELLIPSIS**, which show that an event is related to another event via a shared participant (*Harry filled out the form and _ mailed it to the customs office*).

The three steps of particularizing, structuring, and cohering result in a grammatical structure that evokes a reconstituted version of the original unique whole. These six steps in verbalization are not necessarily processed sequentially or independently. The steps in the verbalization process are dependent on the grammatical resources available in the language, which constrain the possibilities available to the speaker. For example, when a speaker takes a subchunk and extracts participants from it, there must be a construction available in the language to relate the participants to the predicate, as with *put* in the earlier example. Thus, subchunking must be coordinated with propositionalizing and structuring. Also, the steps may not be overtly expressed by grammatical inflections or particles. For example, *The book fell* does not overtly express the singular number of *book*, or that the event is situated in the real world rather than a nonreal mental space of the speaker.

Finally, the reconstituted experience evoked by the linguistic utterance is not the same as the unique whole with which the speaker began. The cognitive processes humans use in verbalization do not simply carry out one or more of the six steps described. They also conceptualize the experience in different ways, depending on the speaker's choices. These choices range from the subtle difference between describing something as *leaves* or *foliage*, or the more dramatic framing differences between *fetus* and *unborn baby* referred to previously. There are a wide range of conceptualization processes or construal operations that have been identified in language (see, e.g., Langacker 1987; Talmy 2000). The construal operations can be accounted for by processes familiar from cognitive psychology: attention, comparison, perspective, and Gestalt (Croft and Cruse 2004, Chapter 4). These psychological processes are part of the meaning of all linguistic units: words, inflections, and constructions. As a consequence, every utterance presents a complex conceptualization of the original experience that the speaker intends to verbalize for the hearer. The conventionalized conceptualizations embodied in the grammatical resources of a language represent cultural traditions of ways to verbalize experience in the speech community.

VARIATION AND THE USAGE-BASED MODEL

One of the results of recent research on language structure and language use is the focus on the ubiquity of variation in language use, that is, in the verbalization of experience and its phonetic realization. The ubiquity of variation in language use has led to new models of the representation of linguistic knowledge in the mind that incorporate variation as an essential characteristic of language. These models are more developed in phonetics and phonology. The phonological model is described first and then recent proposals to apply it to grammar (syntax and **SEMANTICS**) are examined.

One of the major results of instrumental phonetics is the discovery that phonetic variation in speech is ubiquitous. Variation in the realization of phonemes is found not just across speakers but also in the speech of a single speaker. There are at least two reasons why such variation in the speech signal would exist. Presumably, the level of neuromuscular control over articulatory gestures needed for identical (invariant) productions of a phoneme is beyond a speaker's ability. At least as important, the variation in the speech signal does not prevent successful communication (or not enough of the time to lead to the evolution of even finer neuromuscular control abilities in humans).

There is evidence, moreover, that the mental representation of phonological categories includes the representation of individual tokens of sounds and the words that contain them. Speakers retain knowledge of fine-grained phonetic detail (Bybee 2001; Pierrehumbert 2003). Also, there are many frequency effects on phonological patterns (Bybee 2001). For example, higher-frequency forms tend to have more reduced phonetic realizations of phonemes than lower-frequency forms.

Finally, human beings are extremely good pattern detectors from infancy on into adulthood. Infants are able to develop sensitivity to subtle statistical patterns of the phonetic signals they are exposed to. This type of learning, which occurs without actively attending to the stimulus or an intention to learn is called *implicit learning* (Vihman and Gathercole, unpublished manuscript). It contrasts with *explicit learning*, which takes place under attention from the learner – particularly joint attention between an infant learning language and an adult – and is involved in the formation of categories and symbolic processing. There is neuroscientific evidence that implicit learning is associated with the neocortex and explicit learning with the hippocampus (ibid.).

A number of researchers have proposed a **USAGE-BASED** or **EXEMPLAR** model of phonological representation to account for these patterns (Bybee 2001; Pierrehumbert 2003). In this model, phonological categories are not represented by specific phonetic values for the phoneme in the language, but by a cluster of remembered tokens that form a density distribution over a space of phonetic parameters. The phonetic space represents the phonetic similarities of tokens of the phonological category. This model includes properties of implicit learning (the cluster of individual tokens) and explicit learning (the labeling of the density distribution as representing tokens of, say, /e/ and not /i/). Consolidation of token memories also takes place – individual tokens decay in memory, highly similar tokens are merged, and the distribution of tokens can be restructured – but new tokens are constantly being incorporated into the representation and influencing it.

Marilyn Vihman and S. Kunnari (2006) propose three types of learning for an exemplar model. First, there is an initial implicit learning of statistical regularities of the sensory input. Second, explicit learning of linguistic categories, such as the words that are templates containing the sound segments, takes place. Finally, a second layer of implicit learning of statistical

regularities gives rise to probability distributions for each linguistic phonological and lexical category. The result of this last layer of learning is the exemplar or usage-based model described by Janet Pierrehumbert and Joan Bybee.

The application of the usage-based/exemplar model to grammar is more complex. Most research in this area has compared the range of uses of a particular word or grammatical construction. However, this does not represent the process of language production (that is, verbalization), analogous to the phonetic variation found in the production of phonemes. Studies of parallel verbalizations of particular scenes demonstrate that variation in the verbalization of the same scene by speakers in similar circumstances is ubiquitous, much like the phonetic realization of phonological categories (Croft 2010).

There is also substantial evidence for frequency effects in grammar. For example, English has a grammatical category of auxiliary verb that has distinctive syntax in negation (*I **can't** sing* vs. *I **didn't sing***), questions (***Can** he sing?* vs. ***Did** he **sing**?*). These syntactic patterns are actually a relic of an earlier stage of English when word order was freer; it has survived in the auxiliaries of modern English because of their higher token frequency (Bybee and Thompson 1997), as well as their semantic coherence. Frequency plays a central role in the historical process of *Grammaticalization* (Hopper and Traugott 2003), in which certain constructions develop a "grammatical" function (more precisely, they are recruited to serve the particularizing, structuring, and cohering steps of the verbalization process). Part of the grammaticalization process is that the construction increases in frequency; it therefore undergoes grammatical and phonological changes, such as fixation of word order, loss of syntactic flexibility, and phonetic reduction (Bybee 2003). A well-known example is the recruitment of the *go* + Infinitive construction for the future tense: *She is going (to Sears) to buy a food processor* becomes future *She's going to buy a food processor*, with no possibility of inserting a phrase between *go* and the infinitive, and is finally reduced to *She's gonna buy a food processor.*

Finally, early syntactic acquisition is driven by implicit learning of patterns in the linguistic input (Tomasello 2003). The process of syntactic acquisition is very gradual and inductive, involving an interplay between detection of statistical regularities and the formation of categories that permit productive extension of grammatical constructions. Children occasionally produce overregularization errors, and these are also sensitive to frequency (more frequent forms are more likely to be produced correctly, and less frequent forms are more likely to be subject to regularization).

A usage-based model of grammatical form and meaning is gradually emerging from this research. An exemplar model of grammatical knowledge would treat linguistic meanings as possessing a frequency distribution of tokens of remembered constructions used for that meaning. Those constructions would be organized in a multidimensional syntactic space organized by structural similarity (e.g., Croft 2001, Chapter 8) and whose dimensions are organized by the function played by the construction in the verbalization process. The meanings of constructions are themselves organized in a conceptual space whose structure can be inferred empirically via cross-linguistic comparison of the meanings expressed by grammatical categories and constructions. The typological approach to grammar has constructed conceptual spaces for a number of semantic domains using techniques such as the semantic map model (see Haspelmath 2003 for a survey of recent studies) and multidimensional scaling (Croft and Poole 2008).

To sum up, the usage-based/exemplar model can be applied to both phonological patterns in words and grammatical structures in constructions. A speaker's knowledge of language is the result of the interplay between two learning processes. One learning process is the tallying of statistical regularities of tokens of words and constructions with a particular phonetic realization, performing a particular communicative act in a specific social interaction. The other is the organization of these tokens into categories and the formation of generalizations that allow the reuse or replication of these grammatical structures to solve future coordination problems in communication.

VARIATION AND CHANGE: AN EVOLUTIONARY APPROACH

The view of language described in the preceding sections roots both language structure and a speaker's linguistic knowledge in the individual acts of linguistic behavior that a speaker has engaged in and will engage in. It is a dynamic view of language in that linguistic behavior is essentially historical: a temporal series of utterances, each one linked to prior utterances as repeated behavior to solve recurrent coordination problems in social interaction. Each member of a speech community has a history of his or her participation in linguistic events, either as speaker or hearer. This history is remembered in the exemplar-based representation of that member's linguistic knowledge, but also consolidated and organized in such a way that each unique experience is broken down and categorized in ways that allow for reuse of words and constructions in future communication events.

Each time a speaker produces an utterance, he or she replicates tokens of linguistic structures – sounds, words, and constructions – based on the remembering of prior tokens of linguistic structures, following the principles of convention and verbalization described earlier. However, the replication process is never perfect: Variation is generated all of the time, as described in the preceding section. The variation generated in the process of language use can be called first-order variation. Variation in replication is the starting point for language change. Language change is an instance of change by replication (rather than inherent change); change by replication is the domain of an evolutionary model of change (Hull 1988; Croft 2000).

Change by replication is a two-step process. The first step is the generation of variation in replication. This requires a replicator and a process of replication by which copies are produced that preserve much of the structure of the original. In biological evolution, the canonical replicator is the gene, and the process of replication takes place in meiosis (which in sexual organisms occurs in sexual reproduction). Copies of the gene are produced, preserving much of the structure of the original gene. Variation is generated by random mutation processes and by recombination in sexual reproduction.

In language, replication occurs in language use. The replicators are tokens of linguistic structures in utterances (called linguemes in Croft 2000). These tokens are instances of linguistic behavior. The process of language change is therefore an example of cultural transmission, governed by principles of evolutionary change. The replication process in language change is governed by the principle of convention. As we have seen in the preceding section, variation is generated in the process of verbalization, including the recombination of linguistic forms. This represents innovation in language change. First-order variation is the source of language change. Experiments in phonological perception and production indicate that "sound change is drawn from a pool of synchronic variation" (the title of Ohala 1989). Indeterminacy in the interpretation of a complex acoustic signal can lead to reanalysis of the phonological categories in that signal. Likewise, it appears that grammatical change is also drawn from a pool of synchronic variation, namely, variation in verbalization. There is an indeterminacy in the understanding of the meaning of a word or construction because we cannot read each other's minds, our knowledge of linguistic conventions differs because we have been exposed to different exemplars, and every situation is unique and can be construed in different ways. This indeterminacy gives rise to variation in verbalization (Croft 2010), and can lead to the reanalysis of the mapping of function into grammatical form (Croft 2000).

The second step of the evolutionary process is the selection of variants. Selection requires an entity other than the replicator, namely, the interactor. The interactor interacts with its environment in such a way that this interaction causes replication to be differential (Hull 1988). In biological evolution, the canonical interactor is the organism. The organism interacts with its environment. In natural selection, some organisms survive to reproduce and therefore replicate their **GENES** while others do not; this process causes differential replication.

In language, selection occurs in language use as well. The interactor is the speaker. The speaker has variant linguistic forms available and chooses one over others based on his or her environment. In language, the most important environmental interaction is the social relationship between speaker and hearer and the social context of the speech event. This is, of course, the realm of sociolinguistics (see, e.g., Labov 2001, and the following section). Selection goes under the name of propagation in language change.

Selection (propagation) is a function of the social value that variants acquire in language use. First-order variation does not have a social value. Socially conditioned variation is second-order variation. Once a variant is propagated in a speech community, it can lead to third-order variation, that is, variation in linguistic conventions across dialects and languages. Linguistic diversity is the result of language change.

The evolutionary model requires a revision to the definition of language offered near the beginning of this essay. In the evolutionary model, a language is a population of utterances, the result of the employment of linguistic conventions in a speech community. The linguistic system is the result of the ways in which speakers have consolidated the uses of language in which they have participated into their knowledge of the conventions of the speech community. Each speaker's systematic knowledge of his or her language is different, because of differences in the range of language use to which each speaker is exposed.

SPEECH COMMUNITIES AND COMMON GROUND

Language in this revised sense is the product of a speech community: the utterances produced by communicative interactions among speakers. A speech community is defined by its social interactions involving language: Members of the speech community communicate with one another, and the community is defined by communicative isolation from other communities. Communicative isolation is relative, of course, and in fact the structure of human speech communities is far more complex than the structure of biological populations.

Two related phenomena serve to define communities: common ground and shared practice. Common ground plays an essential role in defining joint action and convention, both central to understanding the nature of language. Common ground consists of knowledge, beliefs, and attitudes presumed by two or more individuals to be shared between them. Common ground can be divided into two types: personal common ground and communal common ground (Clark 1996, Chapter 4). Personal common ground is shared directly in face-to-face interaction by the persons. Personal common ground has two bases. The first is the perceptual basis: We share knowledge of what is in our shared perceptual field. The perceptual basis is provided by virtue of joint attention and salience, as mentioned earlier. A shared basis for common ground has the following properties: The shared basis provides information to the persons involved that the shared basis holds; the shared basis indicates to each person that it provides information to every person that the shared basis holds; and the shared basis indicates the proposition in the common ground (Clark 1996, 94). A basis for common ground varies in how well it is justified; hence, we may not always be certain of what is common ground or not.

The second basis for personal common ground is a discourse basis. When I report on situations I have experienced to you in conversation, and vice versa, these become part of our personal common ground. Although we did not experience them perceptually together, we did experience the reporting of them linguistically together. The discourse basis thus involves joint attention (to the linguistic signal), as well as the common ground of a shared language. The discourse basis and the perceptual basis both require direct interaction by the interlocutors. They correspond to social networks, which are instrumental in language maintenance and change (Milroy 1987).

The other type of common ground is communal common ground. Communal common ground is shared by virtue of common community membership. A person can establish common ground with a stranger if they both belong to a common community (e.g., Americans, linguists, etc.). Some communities are quite specialized while other communities are very broad and even all-encompassing, such as the community of human beings in this world, which gives rise to the possibility of communication in the first place.

Clark argues that the basis of communal common ground is shared expertise. Étienne Wenger, on the other hand, defines communities of practice in terms of shared practice: Individuals engage in joint actions together, and this gives them common ground and creates a community (Wenger 1998). Wenger's definition of a community of practice, therefore, requires face-to-face interaction, like personal common ground. However, shared practice can be passed on as new members enter the community and share practice with remaining current members. This is cultural transmission and can lead to individuals being members of the same community through a history of shared practice, even if they do not interact directly with every other member of the community.

Since communities are defined by shared practice, and human beings engage in a great variety of joint actions with different groups of people, the community structure of human society is very complex. Every society is made up of multiple communities. Each person in the society is a member of multiple communities, depending on the range of shared activities he or she engages in. The different communities have only partially overlapping memberships.

As a consequence, the structure of a language is equally complex. A linguistic structure – a pronunciation, a word, a construction – is associated with a particular community, or set of communities, in a society. A pronunciation is recognized as an accent characteristic of a particular community. Words will have different meanings in different communities (e.g., *subject* is a grammatical relation for linguists but a person in an experiment for psychologists). The same concept will have different forms in different communities (e.g., *Zinfandel* for the general layperson, *Zin* to a wine aficionado). Thus, a linguistic convention is not just a symbol – a pairing of form and meaning – but includes a third part, the community in which it is conventional. This is part (iii) of the definition of convention given in an earlier section. Finally, each individual has a linguistic repertoire that reflects his or her knowledge and exposure to the communities in which he or she acts.

The choice of a linguistic form on the part of a speaker is an act of identification with the community that uses it. This is the chief mechanism for selection (propagation) in language change: Ihe propagation of variants reflects the dynamics of social change. More recent work in sociolinguistics has argued that linguistic acts of social **IDENTITY** are not always passive: Individuals institute linguistic conventions to construct an identity as well as to adopt one (Eckert 2000).

LANGUAGE DIVERSITY AND ITS ENDANGERMENT

Variation in language can lead to language change if it is propagated through a speech community. Social processes over human history have led to the enormous linguistic diversity we find today – a diversity that newer social processes also threaten. The basic social process giving rise to linguistic diversity is the expansion and separation of populations into distinct societies. As groups of people divide for whatever reason, they become communicatively isolated, and the common language that they once spoke changes in different directions, leading to distinct dialects and eventually to mutually unintelligible languages.

This ubiquitous demographic process is reflected in the family trees of languages that have been constructed by linguists working on genetic classification. These family trees allow for the possibility of reconstructing not just protolanguages but also the underlying social processes that are traced in them. Even sociolinguistic situations that obscure family trees leave linguistic evidence of other social processes. Extensive borrowing indicates a period of intensive social contact. Difficulty in separating branches of a linguistic family tree indicates an expansion through a new area but continued low-level contact between the former dialects. These can be seen in the dialect continua found in much of Europe, where the Romance, Germanic, and Slavic peoples expanded over a mostly continuous terrain (Chambers and Trudgill 1998). Shared typological (structural) traits may be due to intimate contact between languages with continued language maintenance, or to a major language shift by a social group, resulting in a large proportion of non-native speakers at one point in a language's history.

The spread of human beings across the globe led to the creation of a huge number of distinct societies whose languages diverged. The number of distinct languages that have survived until the beginning of the twenty-first century is about 6,000. Most linguists generally accept the hypothesis that modern human language evolved just once in human history, probably no later than 70,000 to 100,000 years ago. So in principle, all modern human languages may have originated in a single common ancestor. Tracing back the actual lineages of contemporary languages deep into human prehistory appears to be extremely difficult, if not impossible. Nevertheless, there is no doubt that contemporary linguistic diversity is extremely ancient in human history. What we can discover about linguistic history by the comparison of existing languages can potentially shed important light on human history and prehistory.

There are linguistic descriptions of a small proportion of existing human languages, though descriptive work has increased and the overall quality of descriptions has improved dramatically, thanks to advances in linguistic science throughout the twentieth century. It would be safe to say that the diversity of linguistic structure, and how that structure is manifested in phonetic reality on the one hand and in the expression of meaning on the other, is truly remarkable and often unexpected. Many proposed universals of language have had to be revised or even abandoned as a consequence, although systematic analysis of existing linguistic descriptions by typologists have revealed many other language universals that appear to be valid. Linguistic diversity has revealed alternative ways of conceptualizing experience in other societies, as well as alternative methods of learning and alternative means for communication for the accomplishment of joint actions.

But the single most important fact about the diversity of human language is that it is severely endangered. Of the 6,000 different languages extant today, 5,000 are spoken by fewer than 100,000 people. The median number of speakers for a language is only 6,000 (Crystal 2000). Many languages are no longer spoken by children in the community, and therefore will go extinct in another generation. The loss for the science

of language, and more generally for our understanding of human history, human thought, and human social behavior, is immense. But the loss is at least as great for the speakers themselves. Language use is a mode of social identity, not just in terms of identifying with a speech community but as the vehicle of cultural transmission. The loss of languages, like other linguistic phenomena, is a reflection of social processes. The most common social processes leading to language loss are disruption, dislocation, or destruction of the society (language loss rarely occurs via genocide of its speakers). The enormous consequences of language loss has led to a shift in linguistic fieldwork from mere language description and documentation to language revitalization in collaboration with members of the speech community. But reversing language shift ultimately requires a change in the social status of the speech community in the local and global socioeconomic system.

SUMMARY

The scientific study of language in its pragmatic, cognitive, and social context beginning in the latter half of the twentieth century is converging on a new perspective on language in the twenty-first century. Linguistic conventions coordinate communication, which in turn coordinates joint actions. The fragility of social interaction by individuals leads to creativity, variation, and dynamism in the verbalization and vocalization of language. Individual linguistic knowledge (the linguistic system) reflects the remembered history of language use and mediates processes of language change. The continually changing structure of society, defined by common ground emerging from shared practices (joint actions), guides the evolution of linguistic conventions throughout its history. Human history in turn has spawned tremendous linguistic diversity, which reflects the diversity of human social and cognitive capacity. But the unchecked operation of contemporary social forces is leading to the destruction of speech communities and the mass extinction of human languages today.

WORKS CITED AND SUGGESTIONS FOR FURTHER READING

Austin, J. L. 1962. *How to Do Things with Words*. Cambridge: Harvard University Press.

Bolinger, Dwight. 1980. *Language, the Loaded Weapon*. London: Longmans.

Bratman, Michael. 1992. "Shared cooperative activity." *Philosophical Review* **101**: 327–41.

Bybee, Joan L. 2001. *Phonology and Language Use*. Cambridge Cambridge University Press.

———. 2003. "Mechanisms of change in grammaticalization: The role of frequency." In *Handbook of Historical Linguistics*, ed. Brian Joseph and Richard Janda, 602–23. Oxford: Blackwell.

Bybee, Joan L., and Sandra A. Thompson. 1997. "Three frequency effects in syntax." *In Proceedings of the 23rd Annual Meeting of the Berkeley Linguistics Society*, ed. Matthew L. Juge and Jeri O. Moxley, 378–88. Berkeley: Berkeley Linguistics Society.

Chafe, Wallace. 1977. "The recall and verbalization of past experience." In *Current Issues in Linguistic Theory*, ed. Peter Cole, 215–46. Bloomington: Indiana University Press.

Chambers, J. K., and Peter Trudgill. 1998. *Dialectology*. 2d ed. Cambridge: Cambridge University Press.

Clark, Herbert H. 1992. *Arenas of Language Use*. Chicago and Stanford: University of Chicago Press and the Center for the Study of Language and Information.

———. 1996. *Using Language*. Cambridge: Cambridge University Press.

Clark, Herbert H.. 1999. "On the origins of conversation." *Verbum* **21**: 147–61.

Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*. 2d ed. Chicago: University of Chicago Press.

Croft, William. 2000. *Explaining language change: An evolutionary approach*. Harlow, Essex: Longman.

———. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

———. 2003. *Typology and Universals*. 2d ed. Cambridge: Cambridge University Press.

———. 2007. "The origins of grammar in the verbalization of experience." *Cognitive Linguistics* **18**: 339–82.

———. 2010. "The origins of grammaticalization in the verbalization of experience." *Linguistics* **48**: 1–48.

Croft, William, and D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Croft, William, and Keith T. Poole. 2008. "Inferring universals from grammatical variation: Multidimensional scaling for typological analysis." *Theoretical Linguistics* **34**: 1–37.

Crystal, David. 2000. *Language Death*. Cambridge: Cambridge University Press.

Eckert, Penelope. 2000. *Linguistic Variation as Social Practice: The Linguistic Construction of Identity in Belten High*. Oxford: Blackwell.

Grice, H. Paul. [1948] 1989. "Meaning." In *Studies in the Way of Words*, 213–23. Cambridge: Harvard University Press.

Haspelmath, Martin. 2003. "The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison." In *The New Psychology of Language*. Vol. 2. Ed. Michael Tomasello, 211–42. Mahwah, NJ: Lawrence Erlbaum Associates.

Hockett, Charles F. 1960. "The origin of speech." *Scientific American* **203**: 88–96.

Hopper, Paul, and Elizabeth Traugott. 2003. *Grammaticalization*. 2d ed. Cambridge: Cambridge University Press.

Hull, David L. 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press.

Labov, William. 2001. *Principles of Linguistic Change*. Vol. 2. *Social Factors*. Oxford: Blackwell.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar. Vol. 1. Theoretical Prerequisites*. Stanford: Stanford University Press.

Lewis, David. 1969. *Convention*. Cambridge, MA: MIT Press.

Milroy, Lesley. 1987. *Language and Social Networks*. 2d ed. Oxford: Basil Blackwell.

Ohala, John. 1989. "Sound change is drawn from a pool of synchronic variation." In *Language Change: Contributions to the Study of its Causes*, ed. Leiv Egil Breivik and Ernst Håkon Jahr, 173–98. Berlin: Mouton de Gruyter.

Peirce, Charles Sanders. 1932. "Ground, object and interpretant." In *Collected Papers of Charles Sanders Peirce*. Vol. 2: *Elements of Logic*, ed. Charles Hartshorne and Paul Weiss, 134–55. Cambridge: Harvard University Press.

Pierrehumbert, Janet B. 2003. "Probabilistic phonology: discrimination and robustness." In *Probabilistic Linguistics*, ed. Rens Bod, Jennifer Hay, and Stefanie Jannedy, 177–228. Cambridge, MA: MIT Press.

Saussure, Ferdinand de. [1916] 1966. *Cours de linguistique générale*. Ed. Ch. Bally and A. Sechehaye. (*Course in General Linguistics*. Trans. Wade Baskin. New York: McGraw-Hill.)

Searle, John R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.