

Human Cognition and the Elaboration of Events: Some Universal Conceptual Categories

Suzanne Kemmer

Rice University and

Max Planck Institute for Evolutionary Anthropology

One of the burning questions in cognitive science is how human beings conceptualize the world around them. How do we categorize experience, such that we can break it down for purposes of conceptual manipulation (i.e., thought) and communication? To what extent do all humans categorize experience in the same way? This leads to the second question: What conceptual categories are present in human cognition in general, as opposed to categories specific to individual languages, or that distinguish us from our primate relatives?

These questions have been approached from many different angles and with various types of methodology, including, for example, studies of infant cognition (Mandler, 1992, 2000), language acquisition (Bowerman, 1996; Slobin, 1985, 2000; Tomasello, 1992, 2000), and comparative primate cognition (Savage-Rumbaugh, Shanker, & Taylor, 1998; Tomasello & Call, 1997).

Two fields that have approached these questions using linguistic evidence from adult language are cognitive linguistics and language typology. Cognitive linguistics investigates how lexical and grammatical form express semantic content, relating such linguistic analyses to independently attested cognitive capacities and processes (Langacker, 1987; Talmy, 2000). Language typology also studies the relation of linguistic form and meaning, but specifically via the investigation of the range and the limits of

crosslinguistic variation via broad-based samples of the world's languages (Comrie, 1981; Greenberg, 1978).

In this chapter I bring together the latter two kinds of investigation in an approach that can be called *cognitive typology*. The main claim in this research is the following: Recurrent typological patterns reveal the distinctness of a number of basic contrasting types of events to which human beings are sensitive. These categories are not purely perceptual, as they are not directly dependent on perceptual information; rather, they are conceptual categories, used in the chunking and organization of conceptual information for purposes of formulating, manipulating, and communicating thought. The categories I focus on are two basic categories of transitivity of the clause, as well as two other related categories, the reflexive and middle. I show that these four categories together define a restricted conceptual space that constrains the possible types of grammatical systems available for the expression of basic kinds of events. In the process I identify an important conceptual parameter, the degree of elaboration of events, which is not only fundamental for this semantic domain but has general ramifications for human conceptualization and language.

THE COGNITIVE-TYPOLOGICAL APPROACH

The approach of cognitively typology is to observe cross-linguistically recurrent patterns of linguistic expression, particularly grammatical expression. That is, we look at how languages systematically make distinctions in form to express differences in meaning, or, equally systematically, fail to make such distinctions. Thus we can find patterns in the ways that particular forms of expression are used in human languages to convey particular kinds of events or situations.

The idea is that if many languages are found to systematically distinguish between two similar meanings by means of a difference in grammatical coding, then such a difference is cognitively significant; the more widespread the differentiation is cross-linguistically, the more likely it is that there is a universal human propensity to pay attention to such a distinction. Conversely, if languages are recurrently found to subsume two meanings under a single form of expression, then this potential for lack of differentiation is also significant: it suggests that the meanings are cognitively closely related.¹

For example, looking at how the notion of possession is treated in the languages of the world, we find that in many languages, in order to

¹The first proponent and explicator of this methodology and its theoretical basis that I am aware of is Charles Ferguson in his typological study of case (Ferguson, 1970).

say that a person possesses something, one says literally that the thing possessed is "at" or "with" or "in the hand of," or "at the house of" the possessor (Heine, 1997). For example, in So, a language of East Africa, the predicative possessive construction is formed as in (1):

- (1) So
 nek Auca eo-a kus-in
 Neg Auca home-LOC skin-PL
 'Auca has no clothes.' (lit.: 'Skins are not at Auca's home.')
 (Heine, 1997, p. 92, cited from Carlin, 1993, p. 68)

In this language, the construction used for expressing a possessive relation has the same structure as that used for talking about the actual location of an object or person. This formal relation between possession and location is extremely widespread in human languages.

In other languages, possession is treated formally differently from predication of location. In English, we say *I have a book* rather than **the book is at/by/with me*, and similarly in many of the languages of Europe. In Guaraní, a South American language, there is a possessive construction that simply links an optional possessor subject with a possessive noun phrase, as in (2):

- (2) Guaraní
 (Che) che-ajaka.
 (1) my-basket
 'I have a basket.' (Velazquez, 1996, p. 69)

Locational expressions in this language, in contrast, do not make use of possessive markers, but instead generally require adpositional phrases spelling out the location of an object.

Examining the relations of expressions for location and for possession cross-linguistically leads to the conclusion that these categories are related cognitively, and further, that of the two, location is the more fundamental category. Possessive constructions are often transparently locational in form and meaning, or else they are at least historically derived from locational expressions, whereas the opposite relation is rare or nonexistent. Given that we can compare and analyze the relation between the two meanings, and show a plausible link between them, it makes sense to hypothesize that the two categories frequently share the same forms of expression precisely because they are notions that are seen to be similar by humans. At the same time, the fact that some languages treat these same two semantic categories quite differently shows they are conceptually differentiable. In fact, the patterns of relation found suggest that each of these categories constitutes a separate conceptual archetype that can attract its own formal marking pattern; and that both are related to a third category, the predication of

existence, which is also frequently based on locative expression (Clark, 1978).²

Using this basic methodology of comparing form-meaning relationships in the languages of the world for many different grammatical categories allows us to examine which kinds of situations human beings are likeliest to group together, and which they are likeliest to distinguish. In this way we can gain an insight into what kinds and properties of things and situations are significant for the human mind, and hence made the basis of conceptual categories. Further, it allows us, through semantic analysis, to work out the precise relations between the categories distinguished. The result of such an investigation is a network of relations among categories, a kind of "semantic map" of the conceptual domains investigated (Kemmer, 1993a).

The question arises why grammatical categories in particular should be examined, rather than, say, comparing lexical concepts across languages. And why focus on the structure of the clause and its components, rather than other linguistic units? I consider each of these issues in turn.

Typologists focus on grammar, that is, the categories associated with morphology and syntax, because the categories of grammar found in the languages of the world appear to be highly constrained in their meaning or function, compared to the meanings of ordinary lexical items, suggesting some cognitive limitations on the categories of grammar (cf. Talmy, 1988). Thus, for example, when it comes to encoding a function like tense, which (summarized simply) expresses the possible temporal relations between a described event and the moment of speech, there is a relatively small inventory of distinctions that tense morphemes or constructions are found to encode, compared with the infinite number of distinctions that are logically possible (Comrie, 1985). Similarly, for the case of locative and possessive predications considered earlier, which appear to represent universal functional categories in human language, there is only a rather limited number of types of such constructions found (Heine, 1997).

The meanings of ordinary lexical items, on the other hand, by no means fall into such small ranges of possible meanings. Unlike grammatical elements, ordinary words are to all appearances open-ended in regard to the concepts they can encode, and these concepts vary much more dramatically across languages as to how they are expressed (if at all). Lexical words typically convey rich and specific meanings that relate fairly directly to the physical, social, and cultural worlds of the speakers that use them. For example, a language might have words for particular recognized cultural concepts and

artifacts like "the stillness of dawn" or "manic beer strainer," for which other languages lack words. Sometimes such lexical items are paraphrasable by longer expressions in other languages, as in the examples just given, but quite often capturing the relevant concept in another language in even a minimally adequate way is quite difficult, especially if the cultures are very different. Even languages spoken by groups that are broadly socially and culturally similar (such as the linguistic populations of Europe) present innumerable cases of lack of lexical correspondences and of culture-specific lexical concepts, as any bilingual or translator knows.

One might think that perhaps comparing words expressing universally shared human experiences would reveal at least some universal human lexical categories. However, we run into difficulties even here. Words for apparently universal cultural experiences or entities such as "mother," "father," "hand," "eat," "sleep," and "die" certainly do share some conceptual content that presumably corresponds to aspects of universal human experience; yet on closer inspection the meanings of the words compared turn out to differ considerably from language to language. For example, even words meaning "mother" or "child" do not uniformly point to the same class of entities across languages. In some languages the same word is used for a biological mother of a particular child as well as for the sisters of that woman. Some languages have two words for "child," depending on whether the person described is being referred to as someone's offspring, or simply as a non-adult human. Lexical items in languages essentially represent the conventional cultural categories of entities and relations to which their speakers find it useful to make frequent reference. The vocabularies of languages vary from one another precisely to the extent that their cultures do.³

Grammatical meanings, on the other hand, are relatively abstract, sometimes extremely so, and are in general much less obviously relatable to cultural specificities. They are simply much more comparable across languages than the meanings of lexical items. Studying the more highly constrained systems of grammatical categories has allowed a good deal of progress to be made in identifying the range of possible conceptual categories expressed in grammar (cf. e.g., Talmy, 1988). Because these categories are in general abstract and hence not directly "given" in perceptual or other experience, and they recur in a relatively uniform way across languages, we must conclude that such categories are highly significant from a human, cognitive standpoint.⁴

²Clark (1978) used a diverse sample of 30 languages to demonstrate the cross-linguistic marking patterns linking locative, existential, and two predicative possessive types, possessor-topic and possessed-topic constructions. She did not speak of conceptual archetypes, or relate the semantic categories in significant detail, but her study foreshadows later cognitive-tyological theory in many respects.

³In a fascinating research program, Wierzbicka (1992, *Inter alia*) has developed a "natural semantic metalanguage" that involves breaking down the meaning of lexical items (as well as grammatical categories) into a set of universal semantic primitives. Differences among languages are seen in this program as being a matter of differences in the conventional combinations of the primitives. Although the analyses produced within this program are very insightful, identifying the set of primitives needed for the vast and open-ended range of lexical concepts found in human languages has proven to be an elusive task. It is at any rate clear that a great deal more generally accepted progress has been made on the grammatical front.

The second question raised is why the linguistic structures compared across languages by typologists are generally clause-level structures and their components, rather than larger units such as the structure of texts or extended chunks of discourse. The reason for this level of focus is that the clause is the linguistic unit that most closely corresponds to the human conceptualization of simple events, that is, particular conceptualizations of situations or occurrences that are formulated in language in the course of speaking and understanding.

There is, it must be emphasized, no uniformity among languages as to precisely how a non-linguistic, perceived situation will be broken down into clauses for linguistic expression; in fact, there is no such uniformity even among speakers of the same language, given that languages afford their speakers many choices. Nevertheless, there is a relation between perceived situations and their expression in clauses; this relation is mediated by the conceptualization of *events*, which fall into a relatively small number of types. An event type can be thought of as a unit of conceptual structure that allows human beings to quickly and conveniently structure complex, temporally extended experience for purposes of thought and communication.

There appears to be a common core of simple conceptual event types that can be expressed in single clauses across languages, and that reflect certain characteristics of the external world, our perceptual facilities and universal human experience.⁵ These types can be fruitfully analyzed and compared; and one can generalize over them to arrive at a description of the basic structure of a simple event (as is done in the following paragraph). Against the backdrop of this common core, languages show considerable variation as to how much and what types of information they can put into a clause, which aspects of the information communicated by the clause are highlighted versus backgrounded, and even the extent to which a conceptualization is integrated into a single conceived event (single eventhood, like single-clause status, is a matter of degree). And, as we see in the course of this chapter, sometimes a single language will provide the resources for

⁵It must be mentioned that lexicon versus grammar is not a dichotomy, but a continuum. More general and schematic lexical items tend to be recruited for expressing grammatical concepts and functions and gradually become less lexical and more grammatical in form and content over time, a process called grammaticalization (see, e.g., Bybee, Pagliuca, & Perkins, 1994; Heine, 1997; Hopper & Traugott, 1993). The study of grammaticalization is part and parcel of the study of linguistic typology.

This assumption, although not often stated, underlies much cognitive linguistic research; see, for example, the work collected in Talmy (2000). My particular formulation of the relation between the world, event conceptualizations, and clauses is taken essentially from Pawley (1987), who compared English and Kalam, languages of radically different structural type. He found that the two languages differ greatly in the set of conceptual situations that can be expressed as a single clause, but that there is a shared set of "more or less isomorphic" conceptual situations that can be so expressed in both languages. This finding accords with my own experience in typological research, and I take such overlaps to be indicative of common conceptual event structures in human language and their natural correlation with simple clause structure.

speakers to systematically vary the way a given situation is viewed, such that the internal structure of an event can be brought into greater or lesser focus. Thus, event structure is to a certain extent language-specific, even utterance-specific. Nevertheless, all such variation consists essentially of conventional modifications, along certain specifiable dimensions, of the same basic conceptual structures: the simple basic event types that can be expressed in simple clauses in the languages of the world.

Now we come to the basic semantic structure of a simple event. Events are composed of *participants* (the entities involved in situations) and the *relations* among those participants. For example, consider the meaning of a sentence in the form of a simple clause, like *Have some candy*. Here one participant (the unexpressed "you" in the situation) is in a particular relation, in this case a potential possession and consumption relation, with another participant, "some candy." Event structure refers to the overall configuration of participants in an event and their relations: how many participants are there, and what roles do they play in the event in terms of initiation (or lack thereof) of action, impingement on one another, other kinds of relations to one another (e.g., participant roles of various types), and the temporal properties of the predicate relations.

In the following section I introduce a grammatical distinction that is extremely well attested cross-linguistically, and which, I claim, corresponds to fundamental cognitive categories in the form of two basic event structure types that humans utilize in organizing and expressing thought.

TWO BASIC EVENT STRUCTURES

It is well known that languages of the world typically make a formal distinction between transitive and intransitive clauses. At a minimum, the distinction between the two clause types includes the fact that a transitive clause typically has two noun phrases (or other signals of referential participants, such as agreement markers), whereas an intransitive clause has one.

But in addition, many languages have ways of distinguishing the two clause types either distributionally and/or with distinctive overt marking. A common distributional distinction is one in which the verbs of a language divide into classes defined by their possibility of use in transitive or intransitive clauses. Such a division is made with varying degrees of sharpness in a great many languages. Further, there is a widespread occurrence in languages of overt markers that can be added to a verb of one class (i.e., a verb usually or always found in one of the two clause types) to make it usable in the other clause type. Such transitivity and detransitivizing morphology is extremely common in languages, and exists even where the distinction between transitive and intransitive verbs is not

a sharp one (i.e., even in languages where there are many verbs that can be used in clauses of either type).

We might at first take the widespread existence of a formal transitivity distinction to be a reflection of the structure of the world: Some events involve two entities, one acting on the other (as in *I ate some cake*), whereas others simply involve one (as in *She ran away*).

However, it soon becomes evident when examining languages that whether a given verb can appear in a transitive or intransitive clause is not strictly predictable from some language-independent notion of how many entities are involved in a particular kind of event. The inventories of intransitive versus transitive verbs found across languages are far from lining up neatly according to whether a given verb meaning intuitively has one or two participants associated with it. Sometimes a verb that is intransitive in one language most closely corresponds to a transitive verb in another, or vice versa. For example, the English verb *go*, an intransitive verb of motion, can occur with various prepositional elements (as in *go around*, *go across*, etc.) to describe motion with respect to some landmark. To describe essentially the same scenes in Luo, speakers must choose from a set of transitive verbs, including *yoro* "to go across," *luoro* "to go around," *dorojo* "to go in," where the landmark is expressed as a direct object and there is no preposition. Even in the same language, we find different ways of expressing the same basic idea, one transitive, one intransitive: English, after all, does have transitive verbs like *cross* and *enter* alternating with the intransitive *go+PREPOSITION* structures.

Because every language, whether overtly or distributionally, shows some difference between transitive and intransitive structures, and because there are myriad and widespread formal manifestations of such a difference, we must conclude there is something cognitively significant about this difference. Despite the lack of a predictable correlation within or across languages between the idea of a particular action and a particular transitivity structure, there is nonetheless a set of strong regularities that lead us to an understanding of transitivity and its function in language.

For example, it is predictable that the transitive clause structure in any given language will be the structure most typically used with verbs of physical contact or force in which one entity volitionally acts on another, like English *hit punch, beat*, and so forth. And it is also predictable that an intransitive structure will be used with verbs of animate entities moving through space, such as *go run, swim* and similar. Many other types of verbs may fit into each of these structures; but cross-linguistically, these two classes of verbs show the most consistent association with simple transitive and intransitive structures respectively. Further, although some crossover use of each of these verb classes with the other structure is possible, it is precisely with such uses that one finds evidence of a less natural "ft." For

example, the contact verbs in English resist placement in an intransitive structure (*?I hit*); such usage is restricted to particular kinds of participants and interpretations (as in *the bomb hit*) and hence, is far less frequent than the transitive use of such verbs. Similarly, although common English verbs of motion through space might be used in a transitive structure, the direct object is likely to be of a very restricted type, such as a unit of distance (*run a mile*), or else the interpretation is not simple motion, but caused motion and/or metaphorical motion (*run a business*).

Similar kinds of observations can be repeated across languages in various forms. In some languages, one can or must use special marking, that is, transitivity or intransitivity morphology on the verb, to use a verb of one of these classes in the noncorresponding clause structure. In others, the difference is just a matter of distributional frequency: A given motion verb simply occurs more often in an intransitive than a transitive structure; and/or, the most frequently used and semantically general motion verbs occur as intransitives rather than transitives. Taken as a whole, the relation of particular verb types and particular clause structures is not random, but patterned.

These facts support the idea that basic transitive and intransitive clause structures are grammatical manifestations of two important conceptual archetypes: two-participant and one-participant events. Languages differ as to precisely which verbal actions are typically expressed by means of one structure or the other, and moreover within individual languages it is often possible to describe the same occurrence in terms of one or the other structure. But this variation is constrained; certain kinds of actions have a natural affinity for one or the other construction.

Moreover, these two archetypes are experientially based: They correlate with major kinds of "scenes," or types of conceived situations basic to human experience (cf. Goldberg, 1998). The basic transitive structure lines up with humans' propensity to pay attention to actions involving an animate entity volitionally acting on a second entity, and exerting physical force on it that leads to contact. This type of event is what has been termed the *prototypical transitive event* (Givón, 1984) or prototypical two-participant event. The basic intransitive structure corresponds to our conception of situations of motion of animate entities, an event structure called the *prototypical intransitive event*⁶ or prototypical one-participant event. Each of these prototype event structures represents what is called a cognitive model (Lakoff, 1987), a rich conceptual structuring of experience that

⁶There are actually four basic intransitive clause structures in human language, identified in an important cross-linguistic study by Stassen (1999). They include, in addition to the intransitive of action (prototypically animate motion) discussed here, clausal predications of location/existence, property/state, and class membership. Each corresponds to its own experiential conceptual scene. Only the intransitive of action is associated with lexical content verbs, rather than grammatical, relatively contentless verbs or with zero expression.

allows us to quickly and effortlessly categorize complex information and reason on the basis of such categorizations.

These two important conceptual archetypes, although corresponding broadly to linguistic units (transitive and intransitive clause structures), clearly go beyond language to deeper levels of conceptualization. Mandler (1992) found evidence of these basic event structures already in prelinguistic infants. Among other evidence, Slobin (1985) found that children acquiring different languages early on showed evidence of having developed a category used consistently in the expression of situations involving asymmetrical physical force—even where the categories of the adult languages differed from one another and from the categories the children created.

Thus, the linguistic structures are based on, or as we might put it, grounded in the conceptual archetypes. Both language and conceptualization are flexible enough to allow alternative conceptualizations and forms of expression, a property of the mind that will be of significance in the discussion of the reflexive and middle event types in the following section. But as we will see, conceptualizations that are in some sense less typical or less in line with our most entrenched experiences are given some kind of special marking, such as additional formal complexity, and moreover they evince greater cross-linguistic variability in form.

The transitivity prototypes in effect form the endpoints of a scale, rather than representing a bipolar opposition. Situations come in all degrees of similarity to the basic cognitive transitivity prototypes, and it has been well-documented that certain specific kinds of contextual and discourse properties affect the degree of formal assimilation of clauses to the two prototypical clause structures (Hopper & Thompson, 1980; Rice, 1987). Deviations from the conceptual prototypes along certain parameters (e.g., coreference of two participants) lead to formal deviations of various sorts from the two basic clause structures, another manifestation of the special marking referred to earlier. It is worth emphasizing that transitivity is not fundamentally a property of particular linguistic elements such as verbs or even clauses, but is rather an aspect of the cognitive models that structure conceptualizations.

GRAMMATICAL CONSTRUCTIONS IN AN INTERMEDIATE SEMANTIC DOMAIN: REFLEXIVE AND MIDDLE

Transitive constructions and intransitive constructions predicating actions are two of the most basic clause structures in human language. A third clause type that occurs in many languages is the reflexive construction, which describes an event in which one participant acts, not on another entity, but on itself. The following examples illustrate reflexive constructions from languages originating in four different continents (Europe, North America, Australia, and Asia, respectively):

- (3) *Spanish*
 Se vio en el espejo.
 REFL. 3SG.REFL.3SG. in the mirror
 "S/he saw him/herself in the mirror."

- (4) *Nahuatl*
 Tl-to-cacqui.
 1PL-REFL-hear
 "We hear ourselves." (Sullivan & Styles, 1988, p. 34)

- (5) *Ngunu*
 ni-jawulpa-mak-i-é-may?
 3MASC.SG-old man-call-REFL-NEG-PPRS
 "He does not call himself an old man." (Hearth, 1978, p. 286)

- (6) *Turkish*
 Orhan kendini aynada gördü.
 Orhan self mirror saw
 "Orhan saw himself in the mirror." (Underhill, 1976, p. 356)

The reflexive is one of a number of what may be called marked clause types (cf. Langacker, 1991, ch. 8). Marked clause types are, intuitively, clause types that are in some sense less basic than unmarked clause types such as simple transitive and intransitive clauses. This non-basic status is indicated in a number of objectively observable ways, considering linguistic properties such as distribution, frequency, and complexity of form, as well as degree of variability. For example, marked types are relatively restricted as compared to more basic types. Unlike the basic types, they do not necessarily occur in all languages, and even within a language show more restrictions on their occurrence. Functionally, they have very specific semantic/pragmatic functions rather than having a very general or default use; and formally they display more structural restrictions. Hence, in terms of both form and function they are less general in occurrence and as a result are less frequent in discourse than corresponding basic types. In addition, they are structurally more variable across languages and are structurally more complex, or at least never less complex, than the unmarked types.⁷

⁷The identification of units and structures that are more versus less basic in language is the study of what is called markedness: unmarked categories are more basic, and marked ones are less basic. In a pioneering study of markedness in human language, Greenberg (1966) built on earlier classic work by Trubetzkoy and by Jakobson on phonological markedness, extending the theory to morphosyntax. (Croft (1991) has further extended the theory in the area of syntactic constructions.)

Although most of these markedness criteria have not been investigated in detail for reflexive constructions, the reflexive is generally agreed to be a marked clause type. Considering at least structural complexity, in surveys of reflexive constructions such as Faltz (1977) and Kemmer (1993a), reflexive constructions are always at least as structurally complex as their simple transitive counterparts (and, a fortiori, more so than intransitive action clauses, which, lacking objects, are simpler than basic transitive clauses).

Although reflexive marking is not found in every language, Kemmer (1988, 1993a) found that of 31 languages in geographically and genetically diverse languages surveyed, all had a reflexive construction distinct from the simple transitive and intransitive clause structures. With further searching on a wider database of over 100 languages, only a few languages were identified that do not make a systematic distinction between situations in which the actor is the same as the acted-on entity, and situations in which the two are distinct (e.g., Tongan, Tuvaluan, and some other languages of the South Pacific). This near ubiquity shows that the reflexive is an important conceptual category in language; humans find it functionally useful to distinguish situations in which the two participants involved are distinct entities, from those in which they are the same entity.

We can think of the reflexive as a kind of in-between semantic category, falling between prototypical transitive events, on the one hand, and prototypical intransitive events, on the other. Reflexive constructions mark cases in which, like the typical transitive situation, there are two participant roles, but like the intransitive, there is only a single entity involved. In a sense, in a reflexive situation there are both two things involved (two participant roles) and at the same time only one thing (one participant playing both roles in the event).

Looking further at the distribution of reflexive marking on particular kinds of situations in languages allows us to refine this characterization further, and in fact to discover yet another closely related conceptual category in the conceptual realm intermediate to the two opposing poles of transitive and intransitive events.

Reflexive constructions are often used in languages to express situations of a person's acting on their own body, for example, shaving, washing, bathing, or grooming the hair; or changing their body position by moving the body into a sitting, standing, or lying position. We can refer to all such actions, for convenience, as body actions. Examples (7) through (10) show the use of the reflexive marker for body actions in languages native to Europe, North America, Africa, and Asia, respectively.⁸

⁸In Djola and Turkish, the verbal affixes shown are not reflexive markers, because in each language they contrast with another form used productively to signal reflexive semantics. I have glossed these affixes MM for "middle marker," which will be explained later.

- (7) *Spanish*
 Maria se pein-ó.
 Maria 3.REFL. comb-PRET
 "Maria combed her hair." (lit. "combed herself")

- (8) *Nahualt*
 me-tema.
 3SG.REFL.bathe
 "S/he bathes him/herself." (Sullivan & Styles, 1988, p. 34)

- (9) *Djola*
 ni-pss-ɔ-pss-ɔ i-ban.
 1SG-wash-MM-REDUP 1SG-finish
 "I have finished washing." (Sapir, 1965, p. 52)

- (10) *Turkish*
 cocuk giy-in-di.
 child dress-MM-past
 "The child got dressed." (Underhill, 1976, p. 354)

Because in such examples the person is acting on him or herself, just as in the cases exemplified in (3) through (6); and because the same marker is used in many languages (e.g., Spanish and Nahualt) for both kinds of situations, the body action cases in (7) through (10) are often assumed to be semantically identical with reflexive situations. As a result, many grammar writers use body action verbs when giving examples of the reflexive construction in the language they are describing.

However, more in-depth cross-linguistic investigation shows that the two sets of cases are not identical, and considering body action verbs as typical examples of reflexive verbs is mistaken. Rather, body action verbs represent a distinct event type that is related to, yet descriptably different from, reflexive semantics. This semantic category is called the *middle voice*, and its grammatical expression is termed a *middle marker*.

First let us consider some formal evidence for distinguishing a separate grammatical category that body action verbs exemplify. There are a good many languages that formally distinguish body actions from reflexive constructions, that is, they use two different constructions for these types. In the following, the (a) examples represent body action cases, and the (b) examples reflexive constructions. (The gloss MM stands for middle marker.)

- (11) *Luo*
 a. Nyako mak-ore.
 girl dress-MM
 "The girl got dressed."

- b. *Woi her-ore kend-e*
 boy love-MM self-3sg
 "The boy loves himself." (Field data)⁹

(12) *Hungarian*

- a. *Borovál-kor-ot*
 shave-MM-PAST-3SG
 "He shaved." (Haiman, 1983, p. 805)
- b. *Fel-emel-t-e mag-á-t*
 up-lift-PAST-3SG self-his-ACC
 "He lifted himself up." (Haiman, 1983, p. 797)

In these languages, there are two constructions, one used productively with transitive verbs in general to form the reflexive construction, and the other, more restricted, used with body action verbs and a range of other situations that are clearly not reflexive in meaning (some of these will be described later). The difference in usage of these constructions goes along with a systematic difference in form: The productive reflexive marker is cross-linguistically almost always a more phonologically substantial form than the middle marker, and is certainly never less substantial, a fact to which I return later.

This systematic differentiation already suggests that there is something special about body actions that distinguish them from reflexives, despite their similar semantics. But the pattern goes further: In languages having such a formal contrast, there is also a systematic meaning distinction that appears when the two constructions are compared.

English is a language that illustrates the pattern very well. Consider the examples of bodily actions given in (13) and (14).

- (13) a. He quickly shaved and got dressed.
 b. I don't need a barber to shave me—I shave *myself*.
 c. Tammy dressed herself today.
- (14) a. I sat up, looked around, then got up and walked away.
 b. I pulled myself up, looked around, then dragged myself to my feet and staggered away.

The (a) examples illustrate normal uses of body action verbs. In English, these are expressed as intransitive constructions of various sorts; there is no pronominal or other object, so we can think of them as zero-marked. The (b)

⁹The Luo data were collected in my Field Methods class at Rice University (1997–1998). It represents a dialect spoken near Lake Victoria in Kenya, which is in some respects different from the Luo represented in published sources available to me. (Tone is not shown here.)

and (c) examples, on the other hand, contain the productive reflexive form PRONOUN+*self*, an overt, phonologically substantial form that designates the acted-on entity.

In addition to the extra reflexive marking found with these body action verbs, the (b) and (c) examples all express situations associated with special circumstances of some type. Example (13b) involves a contrast between two different potential actors, such that the reflexive pronoun points back to one of them, to the deliberate exclusion of the other. From Example (13c) we are likely to infer some out-of-the-ordinary circumstance or some specific kinds of participants, rather than just anyone: For example, we can readily believe that Tammy is a small child who has not hitherto been able to dress herself, or perhaps a woman who is handicapped and has trouble putting her clothes on each morning. In either case, the person described does not have full control of her limbs—it is as though there is some impediment between Tammy's intentionally acting mind, aiming to get dressed, and her relatively passive body. The limbs are not acting in the way normally expected by able-bodied adults, that is, as a direct extension of the will of the acting person.

A similar situation holds in Example (14b). Rather than the relatively effortless motions described in Example (14a), the motions here are effortful—as though the body is a dead weight, rather than a participatory medium for executing the actions directed by the intentional mind. The verbs used are not typical body action verbs this time, but ordinary transitive verbs usually used for designating force applied to inanimate objects (*pull, drag*).

Parallel distinctions are found in other languages with two contrasting constructions. In Russian, for example, we find the following pair, in which the middle marker *-sja* represents the normal way of expressing body actions in Example (15a), but the reflexive marker *sebjá* is used in the special case of contrast with another potential participant in Example (15b).

- (15) *Russian*
- a. *Ja kazdyj den' mojus'.*¹⁰
 I every day wash-MM
 "I wash every day."
- b. *Ja myl sebjá.*
 I washed self
 "I washed myself" (not someone else) (Haiman, 1983, p. 804)

In (15b), there is a contrast made between potential objects of the washing that the speaker has in mind—another person, versus the speaker him or herself. The result is a special emphasis put on the direct object, expressed as extra stress.

¹⁰The Russian middle marker *-sja* is in standard transliteration orthography spelled *s'* after vowels.

In language after language in which such a pair of constructions is found, a similar semantic differentiation holds. The fuller reflexive construction, when used with body actions, is the unusual case: It denotes a contrast between two potential participants, as in Examples (13b) and (15b); or a separation between the acting and acted-on entities, as in (13c) and (14b), that is not the usual conceptualization for humans carrying out body actions.

These differences in the meaning of the reflexive and middle constructions in body actions lead us to conclude that the fuller, reflexive constructions have an essentially different function from the lighter constructions more usually used for body actions. Moreover, these cases provide us with a clue as to how to characterize the semantics of each of these constructions, in a way that not only captures the cross-linguistic generalizations regarding their contrasting use, but also more generally the distribution of reflexive and middle constructions in all languages.

Let us consider first the range of application of reflexive and middle constructions. Where there are two constructions in a given language, the fuller one is always the general reflexive form, and is thus used productively with transitive verbs to form marked, reflexive forms. This systematic alternation with simple transitive events suggests that the reflexive is considered a special case of such events—the case in which the two participant roles that are evoked in a transitive event happen to be filled by the same entity.

The other form, in contrast, cannot be used across the board with transitive verbs: It is restricted to particular semantic verb classes that include the body action classes exemplified earlier. All such "light" forms can be used with at least some verbs designating typical actions performed on the body. Moreover, they are typically also found with other, non-reflexive classes of verbs: verbs of cognition and emotion, for example, as well as a particular kind of reciprocal action, verbs of spontaneous process, and a number of other recurrent categories described in Kemmer (1993a). The verb roots found with this construction have a number of things in common: for one thing, none are prototypically transitive verbs, that is, verbs whose meanings are across languages associated with transitive constructions. A number of these non-reflexive predicate types, for example, the verbs of cognition, are generally *intransitive* in their usual (non-reflexive-marked) use, rather than transitive. In fact, very often there is no corresponding non-marked form at all for verbs found with the lighter construction, whether body action verbs or not. Such non-alternating forms are called "deponents." Middle-marked deponents are typically found, and are perhaps universal, in languages with middle markers. Some examples of light-marked deponents include Turkish *Is-in* "become warm," Latin *vereo-r* "fear," and Old Icelandic *grœma-st* "turn green." Full reflexive forms, in contrast, are never deponents.

These observations lead us to an explication of the semantics of the two types. Reflexive and body action verbs have in common the fact that they refer to situations in which there is an entity that is in some sense the source or origin of the event, as well as at the same time being an entity affected by the event. This entity plays in both cases what we might call an initiating and an endpoint role in the conception of the event.

Beyond this, the event structures of reflexive and body action verbs diverge. With reflexive semantics, the types of predicates that occur are those with an inherent conception of two roles and two participants in those-roles—that is, prototypical transitive events. The reflexive marks the unexpected case that those two roles are filled by the same individual. The body action situations, on the other hand, are different—they are events in which the initiating and affected entity are *predictably* the same, that is, the sameness is inherent to the semantics of the predicates. There is no expectation that the two will be different entities; only one entity is thought to be involved in such actions in the first place. Thus, the two kinds of events differ in how conceptually distinguished the participants are from one another. In the reflexive, two roles are distinguished that are both played by one individual;¹¹ in the body action type, there is not even the conceptual differentiation of two participant roles. The body action types have, in effect, a partially fused participant conception: The participant is conceptually complex enough to have both an initiating and an endpoint aspect, but not so conceptually differentiated as to have two full-blown participant roles associated with it.

The cases of the body actions unusually expressed with reflexive markers illustrated in (13b), (13c), (14b), and (15b) support this analysis. In the case of contrast (13b and 15b), the linguistic context introduces an idea of two roles, potentially filled by separate participants. When a reflexive marker is used contrastively with a body action, it is consistent with the universal function of reflexive markers—to signal the unexpected identity of the actor and acted-on participant. The contrastive context, however, has effectively forced a differentiation between aspects of an entity that are not usually distinguished.

The analysis holds also for the case of reflexives used with participants lacking full control over the body, as in (13c) and (14b) (also 12b). An impediment or other lack of full control is an indication of a separation

¹¹Tahmy (2000, pp. 460–461) described the English reflexive construction in the realm of mental events (such as *control oneself*, *restrain oneself*) in terms of "the divided self": the controlling aspect of a person pitted against unconscious or involuntary desires. There are some fascinating complexities of mental predicates in interaction with both reflexive and middle marking that require deeper analysis than I am able to give them here; suffice it to say that mental events are quite intermediate semantically and are susceptible to either reflexive or middle construal (cf. Kemmer, 1993a, pp. 127–142).

between acting on and acted-on entities that is unusual with ordinary grooming or other body type actions. Such actions normally have a single participant that is not overtly distinguished into the two aspects of body and mind: The two potentially distinguishable aspects are, except in these unusual kinds of situations where there is a special reason to highlight this duality, conceptually fused into a single, less differentiated entity.

The partial conceptual fusion of participants characteristic of the middle is mirrored by the more minimal, morphologically fused expression conventionally found with body action types, that is, the light forms. Reflexive markers are very often full lexical forms, for example, nouns or pronouns serving as direct objects, whereas the form used in the expression of body actions is typically a bound morpheme, or else simply an intransitive construction of some sort (i.e., it has no overt marking separate from the verb root). This situation is illustrated in Hungarian in (12), in which the middle verbal affix *-kod/-koz-* is contrasted with the full, inflected noun object *mag*. Sometimes, as in Luo (Example 11b), the reflexive form is effectively composite, consisting of the middle marker with the addition of another element, often lexical. In still other cases, both forms are bound verbal affixes, but the reflexive simply has more phonological material than the other form (e.g., Djola *-m* "REFLEXIVE MARKER" vs. *-o* "MIDDLE MARKER"). I am not aware of any case in which the reverse relation in substance holds; the reflexive form is always at least as substantial as the middle form, and in most cases more substantial. The relation of the form and the meaning of middle and reflexive markers is taken up in more detail in the following section.

Body actions, it turns out, represent only the most prototypical type of situation among many that share the property of having lower conceptual differentiation of participants vis-à-vis the reflexive. Other kinds of events have this property, and are in fact expressed with the same light markers as the body action types. This is further evidence that languages treat partial conceptual fusion of participants as a significant property, susceptible to conventionalization as a grammatical category distinct from the reflexive. Light marking contrasting with heavy reflexive marking is found recurrently on, besides body actions, verbs of cognition ("think," "consider," "believe," etc.), emotion ("fear," "be happy"), spontaneous process ("break," "collapse," "melt," "rot"), and on naturally reciprocal actions ("kiss," "touch," "meet"; cf. the later section "Degree of Elaboration of Events"). All of these cases of lower conceptual differentiation of participants can be referred to as middle categories.

Each of these semantically middle types is related to the body action types in specifiable ways, of which I give just one example here. Mental predicates, which include the cognition and emotion type events, are like the body actions in that there is a participant that is seen as in some sense

initiating the event. The participant called for in such events is a person whose mental action the predicate describes, that is, a conceptualizer. This conceptualizer has to provide some mental initiative for events of these types, if only to direct his or her attention to a perceptual stimulus that gives rise to the thought or emotion. At the same time, the conceptualizer is also an affected participant, who is aware of and indeed participating in the thought or emotion experienced. These types of events are more obviously different from reflexives than the body actions, because it is harder to think of the affected entity (the mental experienter) as someone acted on. In the case of body actions we can think of the event as a person acting on their own body, and we can even contrast acting on one's own body with acting on someone else's (as in Examples 13b and 15b). Events of cognition, in contrast, are not so readily decomposable into two such aspects of an individual. This is why such events are cross-linguistically generally intransitive: They are intrinsically one-participant events whose single entity, a conceptualizer, resists conceptual decomposition. Mental events in which such a partial decomposition is in fact made are therefore a subtype of middle semantics, one that happens to evince minimal conceptual differentiation.

Reflexive markers, as we saw previously, primarily signal unexpected coreference between two participants. This is their basic function in all languages, including those that do not differentiate between reflexive and middle constructions. In such languages, the reflexive form extends to cover the body action situations and often the other non-reflexive semantic classes typically found with light verbs as well. For example, the Spanish and Nahuatl constructions illustrated in (3) and (4) are used both as productive reflexives and also for body actions (e.g., 7 and 8) as well as for other middle verb types (cognition, emotion, naturally reciprocal action, spontaneous process, etc.). Yet despite this lack of formal differentiation, the difference between reflexive and the middle semantic types still emerges distributionally—only the middle verb types, including the body actions, exhibit deponent verbs, which, as mentioned earlier, are verbs that lack a transitive counterpart and have middle meaning. Deponents from languages with no formal distinction between reflexive and middle include German *sich nähern* "come close, approach," Spanish *encapzarse* "cloak oneself or put one's hood on," Kanuri *har-t-in* "wash oneself (while partially dressed)" (Hutchinson, 1981), and numerous Guugu Yimidhirr verbs including *daga-dhi* "sit down" and *mira-dhi* "show oneself" (Haviland, 1979, p. 126).

We can conclude from this discussion that the reflexive and the middle are two distinct semantic categories, whose distinctness is shown by the fact that many genetically and geographically diverse languages (Hungarian, Turkish, Djola, English, and Russian, among many others) distinguish

them: and, at the same time, whose close semantic relation can be seen from the propensity of many other languages to subsume them under the same formal expression (these include Spanish, Nahuatl, Guarani, Kanuri, and Guugu Yimidhrr, among a great many others). The closer a given middle use is to the reflexive, in fact, the more likely it is to be given the same formal marking as the reflexive. The body action types are the middle uses that are most similar semantically to the reflexive proper, and indeed they designate the kinds of situations that are most often subsumed under reflexive marking across languages.

A second conclusion we can draw is that the two categories established differ in regard to the degree to which the participants involved in the events are conceptually distinguished from one another. The reflexive distinguishes two participants to the extent that it evokes two distinct participant roles in its conception, which are then signaled as being the same entity. The middle, on the other hand, evokes no such role distinction. There is only one participant, but it is conceptually internally complex, given that reference is made to its inherent initiating and affected aspects.

We have thus identified a conceptual parameter that links and at the same time differentiates the reflexive from the middle. We can call this parameter the *relative distinguishability of participants*. We can also generalize further by noticing that this parameter actually relates all of the categories in the realm of transitivity that we have considered so far. The reflexive, as pointed out in the previous section, falls in between the prototypical transitive and intransitive event types. It is in between in the sense that the entity involved is partially distinguished into two: Unlike the transitive type, which has a full-blown differentiation of two participants and participant roles, the reflexive has one actual entity filling its two participant roles. The middle fuses the participants further, by having a single, internally complex participant. And the prototypical intransitive event type displays the opposite extreme from the transitive: It is characterized by a single and conceptually completely undifferentiated participant.

Figure 3.1 sums up the relation among the four transitivity categories. The categories at the opposite extremes of distinguishability are the transitive and intransitive prototypes; the intermediate ones are the reflexive and middle.

Figure 3.1 effectively represents a semantic map of the type described in the Cognitive-Typological Approach section. The particular set of semantic relations portrayed is based on just one dimension of relationship; it is somewhat analogous to a simple map of a road showing only the sequence of towns lying along it. It is useful in that it shows the relative semantic proximity of the categories along this dimension: the middle is semantically farther away from the transitive and closer to the intransitive prototype. This arrangement generates predictions about which categories are most likely

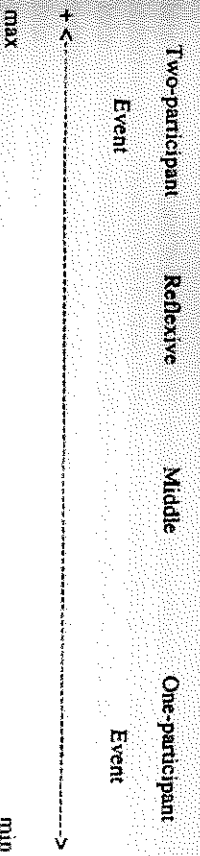


FIG. 3.1. Relation of transitivity categories by degree of participant distinguishability.

to have similar marking in the languages of the world. Non-contiguous categories are predicted to have the same grammatical marking only when the immediately adjacent intermediate categories also share that marking.

A more complex semantic map of the middle and related semantic categories that makes explicit the relations of the various middle subcategories and that takes into account other dimensions of semantic relation is found in Kemmer (1993a, ch. 6).

DEGREE OF ELABORATION OF EVENTS

We can generalize still further when we observe that distinguishability of participants is just one aspect of a more general conceptual parameter. Just as one can distinguish participants to a greater or lesser degree, one can also do the same with the relations among participants, the kinds of events and subevents that, along with the participants, make up the overall event structure.

A good example of the way that languages can distinguish events to different degrees is found with reciprocal constructions, which express situations in which two participants act on each other. In the situation described by *Alice and Ted kissed each other*, Alice and Ted are performing similar actions, but in opposite role configurations: Alice kissing Ted, and Ted kissing Alice. English, like many languages, allows an alternative form of expression for such "mirror-image" actions. In English we can employ verbs with reciprocal semantics intransitively, as in *Alice and Ted kissed*. With the latter expression, it is clear that a single kissing action is portrayed in which both protagonists have engaged. In the alternative, overtly marked reciprocal construction with *each other*, however, we can easily obtain the interpretation that the two kissed each other sequentially, for example, each giving the other a peck on the cheek one after the other. In effect, the two forms of expression provide alternative ways of describing similar situations: one in which there are two separable subevents taking place (temporal separation being one aspect of event separability), and the other in which the event is a relatively undifferentiated whole.

A number of languages, including Russian, Turkish, Hungarian, and Icelandic, have similar differentiations between two reciprocal constructions. As in the case of the body action middle, specific predicates with special inherent semantics are involved; and as with the body actions, a lighter marking is found with these predicates than is found with ordinary reciprocal constructions. The predicates found with light reciprocal marking are those in which the action is very often or expected to be reciprocal between two participants: "argue," "fight," "meet," "wrestle," "embrace," and similar verb meanings, what can be termed naturally reciprocal actions.

Because the light reciprocal marker is in many languages the same as the middle marker, and even when it is not (e.g., in Turkish) it patterns with the middle marker in terms of relative formal substance, and because of the semantic relativity of the light-marked reciprocal category to the middle in terms of distinguishability, it makes sense to consider this special reciprocal a subcategory of middle semantics. The two types of distinguishability (of participants, and of events) can be generalized to an overarching semantic parameter that I call *degree of elaboration of events*: the degree to which the components of an event structure, whether participants or events/subevents, are conceptually distinguished into more fine-grained components.

The notion of degree of elaboration of events, it turns out, is applicable to a wide range of grammatical phenomena, including collective marking on both verbs and nouns (Kemmer, 1993b), and many other kinds of alternating structures involving the degree to which events are conceptually fused versus separated: causatives (Haiman, 1983); complement constructions (Givón, 1980); serial verbs; and noun incorporation constructions (Velazquez, 1996). In all of these cases, languages provide a conventional means for categorizing similar situations as involving, alternatively, elements with more conceptual separation, and those with more conceptual fusion.

This general phenomenon highlights the crucial role of language as a means of providing conventionalized construals, or shared ways of viewing and portraying the situations that are the subject matter of communication (Langacker, 1987). Human language, unlike other forms of animal communication we know about, has the flexibility to provide its users with multiple ways of conceiving and expressing a given situation, adapted to speakers' various discourse and other communicative purposes.

In the case of the degree of elaboration of events, it appears that we are dealing with a capacity that is fundamentally an attentional one. We can, as it were, turn up or turn down the "resolution" that we bring to bear on a given conceptual content, accordingly as our language makes available conventional categories to do so. In the specific case of the voice and transitivity categories we have been discussing, if a language provides the requisite communicative resources, speakers may place a greater focus of

attention on an event's component parts, such as participants or component subevents, and the result is a relatively more elaborated event structure such as a reflexive or an ordinary reciprocal. If, on the other hand, it suits such a speaker's communicative purpose to leave these aspects of the situation relatively undifferentiated, then the situation is categorized as a middle. Languages that happen to lack the formal category of middle categorize such conceptually in-between situations as body actions as either reflexive, or as intransitive, one-participant events. In such cases basically middle semantics is assimilated to a certain degree to the semantics of similar constructions and hence given the form of these other constructions.¹²

It is an open question as to whether there are unambiguously identifiable factors that determine how many categories in this general domain a language will have available (within, of course, the constraints that appear to govern the number and types of categorizations made as identified previously). What makes one language, like Spanish, distinguish the reflexive as a distinct conventional category but not the middle, whereas another language, like Turkish or Russian, separately distinguishes all four of the possible categories of voice and transitivity identified here? It is unlikely, given the odd groupings of languages found with these various categories, that such differences correlate with cultural or social factors.

The best that can be said at this point is that what grammatical categories a language conventionalizes, within the available human linguistic possibilities, is a matter of historical contingency. Every language comprises a complex ecology of available and competing forms, a vast system of linguistic knowledge shared to a certain extent by speakers of the language, yet displaying patterns of heterogeneity across subgroups of those speakers. This complex system absorbs motivated innovations and accommodates itself to gradually changing patterns of linguistic usage, as innovations are adopted and spread through the community. A language will have a dedicated middle voice marker just in case at some point in its history some speakers begin to extend an appropriately similar category, such as the reflexive, to express situations that differ semantically from the reflexive, such as the body actions; and when that marker at some point loses its formal connection with the reflexive marker. Precisely this development has

¹²That middle situations never quite lose their semantic identity as middles, even in languages that have no formal category for the expression of the middle, is seen in a number of ways. For example, there is evidence that middles do not quite behave entirely like reflexives syntactically in languages in which the two are putatively "the same" category (cf. Kemmer, 1993a, pp. 216-218). Further, languages lacking a formal distinction between reflexive and middle show dependents in the same verb categories as those that do have distinct reflexive and middle forms, as noted in the Grammatical Constructions in an Intermediate Domain section.

taken place in many languages in different parts of the world (see Kemmer, 1993a, ch. 5 for descriptions of these developments).

Whether there are any structural, ecological factors within a linguistic system that will influence if and when such an innovation occurs or begins to spread is unknown. To investigate this, we need a typological database of correlations of grammatical properties in the languages of the world, a vast project that has hardly been attempted.

To summarize, the degree of elaboration of events is a broadly significant linguistic and conceptual parameter. It is a subcase of the general human capacity to construe situations in alternate ways, and of the propensity for the grammatical systems of languages to conventionalize specific construals and potentially contrast such construals. Elaboration of events is a specifically attentional phenomenon, relating to the possibility for viewing situations at different levels of detail, or granularity, dependent on relevant communicative and contextual factors. Although the possibilities for grammatical categories referring to elaboration of events in the domain of transitivity and voice are highly constrained, the existence of any structural factors influencing the precise division of such categories in specific languages is unknown.

ICONICITY

We saw in the section on Grammatical Constructions in an Intermediate Domain that in languages having both a reflexive and a middle marker, these two forms align in a predictable way with the two categories they express. The marking associated with the category of reflexive is recurrently associated with greater phonological substance, compared to the form used for the middle, which is phonologically lighter. Haiman (1983) pointed out this basic contrast, and explained it in terms of two functional motivations: iconic and economic motivation.

Economic motivation refers to the tendency, found in communication systems, to minimize formal signals for known or predictable information. An illustration is found in e-mail addresses, which in the United States do not have a country code suffixed to them. The Internet began as a mode of communication within the United States, so the country of destination was predictable and thus did not need to be overtly signaled. When computers outside the United States began to be linked to the Internet, non-U.S. addresses were the "unexpected" case, and country codes began to be added to clarify the destination.

According to Haiman (1983), the light marker is phonologically light because the identity of actor and acted-on participants is predictable, in fact inherent to body action and naturally reciprocal verbs. Thus such verbs

do not require much, if any, extra marking to signal this function. On the other hand, full reflexive and reciprocal forms signal the case of unexpected reflexivity and reciprocity. As we know from the study of markedness, it is the unexpected cases that require extra marking, not only in this semantic domain, but in language and communicative systems more generally.

The other functional motivation described by Haiman (1983) is *iconic motivation*, the tendency to align properties of linguistic form with observably similar properties of linguistic function or meaning, such that form appears to mirror function. In the case of the reflexive/middle contrast, Haiman found that the tendency for reflexives to be free forms, rather than bound verbal affixes, as opposed to the almost inevitably bound middle markers, shows a remarkable iconic correlation with the function of the markers, in terms of the notion of distance: Reflexive forms show more formal distance from the verb root, and at the same time exemplify, in their semantic structure, a greater conceptual separation (and hence distance) between the actor and acted-on entity. The latter, as we saw in the Grammatical Constructions section, are conceptually differentiated into two different participant roles. The middle, the form that is more formally bound up with the verb root, likewise signals an acted-on "entity" that is less conceptually separable from the actor—indeed, from the event itself. A conceptually fused participant displays minimal conceptual distance between its initiating and affected aspects.

Both of these two independently attested motivations, the iconic and the economic, are operative in the case of the middle and reflexive and the two types of reciprocal, as Haiman showed. I would like to go further and say that conceptual distance is just one aspect of what we have identified as degree of elaboration of events; and secondly, that the correlation goes beyond the categories Haiman considered to the broader domain of transitivity introduced in the section on Two Basic Event Structures.

First, let us consider what conceptual dimension is at issue in these categories. Iconic motivation is operating not just in terms of distance, but also of what we might call "weight" or substance. Just as degree of formal distance mirrors conceptual distance, degree of formal weight or substance also corresponds to the degree of conceptual substance associated with the respective semantic categories expressed by the two forms. The more conceptual "material" there is in a given expression, the more formal substance there will be to signal it. Formal substance comprises segmental material and other phonological properties of a marker; morphological substance is measured in terms of degree of morphological/lexical autonomy (degree of boundedness and invariability); and syntactic substance in terms of syntactic autonomy and degree of constituency. Conceptual substance corresponds to the degree of elaboration of an event: how distinct are the participants, from one another and from the event itself (this is degree of

distinguishability of participants); to what extent are the events "pulled apart" into temporally separable subevents (degree of distinguishability of events); and most generally, what and how many things, roles, and relations are identified in the conceptualization the speaker wants to get across or to understand. It is this "pulling apart" of participants, elaborating them into conceptual units in their own right, that gives the effect of adding conceptual distance to an expression.

When we think about it, a general correlation between degree of elaboration of events, and the amount of formal substance used to talk about the events, makes functional sense. We can expect that speakers will reach for more linguistic material to express the events and participants they wish to place in the forefront of consciousness and attention, and will leave unexpressed or minimally expressed those aspects of the scene that are not, for whatever reason, placed in the communicative foreground. If the reason for this attentional backgrounding happens to be that such aspects are already known or predictable, then we have a convergence between iconic and economic motivations.

What is interesting about the particular correlations found in the reflexive and middle domain (and the reciprocal domain) is that they go beyond a general or rough correlation between amount of form and amount of content. In this case, we find that some significant subset of languages has developed a grammatical opposition between minimally contrasting pairs associated with rather fine distinctions in degree of elaboration and, moreover, these precise distinctions are cross-linguistically recurrent. These facts suggest that this particular correlation has a specific function. What this function is, is not obvious, but we can suggest as a first hypothesis that such a strong and precise iconic correlation in the realm of reflexive, reciprocal, and middle semantics perhaps makes the associated conceptual distinctions and their corresponding forms easier to learn and/or process. The correlation might have a partially redundant communicative function such that, in effect, the amount of formal substance of the linguistic form helps to actually signal the amount of conceptual substance comprising the event to be communicated.

The question of the function of iconicity and economy in the learning and processing of language has been barely recognized as an issue in linguistics, let alone explored in any detail. Yet such investigation promises to shed light on one of the most fascinating aspects of human cognition, and one which may link human communication functionally with the communicative systems of other organisms.

SOME RESULTS AND CONCLUSIONS

Having examined the range of possibilities in marking patterns in the semantic domain of reflexive and middle marking, we can draw some interesting conclusions about human conceptual organization.

First, it is significant that human language recurrently provides grammatical distinctions for a conceptual dimension like degree of participant distinguishability. Instead of having a simple one-versus-two-participant distinction in clause structure, which would presumably be grounded in some human cognitive ability to perceive that either one or two things (or groups of things) are carrying out particular roles in an event, languages overwhelmingly tend to add a third category, the reflexive. The reflexive covers cases that are two-participant-like, yet also one-participant-like: A person acting on him or herself is clearly just one entity, but it is one that, by fulfilling two participant roles in the same event, effectively functions like two entities. Then, a smaller number of languages have a fourth grammatical category for an even more subtly "in-between" category, the middle, which is designed to express situations in which there is only one participant, but one that is conceived as in some measure internally complex.

A second noteworthy result is that the scale of transitivity, which had been found independently to be of importance in human language, correlates with the more general conceptual dimension of degree of elaboration of events. The categories of reflexive and middle exist in order to express divergences from canonical or archetypal event types falling at opposite extremes along the scale of semantic transitivity, which correlates directly with a scale of relative degree of elaboration of events. Marking systems in the reflexive/middle domain integrally involve alternative conceptualizations of participant structure and event structure as a whole. Languages may differ as to the morphosyntactic means they have conventionalized for coding such differences in conceptualization, but the variation is highly constrained by the underlying conceptual system.

Thus we have a situation in which (a) humans make a finer-grained set of conceptual distinctions than would be warranted by the sheer perceptual differentiation between one entity versus two; (b) the range of conceptual distinctions is potentially infinite, because once cut loose from perception, all degrees of variation in conceptual event structure along the relevant parameters are logically possible; yet (c) only a maximum of four distinct grammatical categories are found in the languages of the world to express differentiations in degree of participant distinguishability and more generally elaboration of events. The attested variation in languages as to precisely which types of situations are grammatically distinguished in the reflexive/middle realm is thus sharply limited, both by the semantic parameters defining the variation space, and also by the semantic relations between the various categories. Categories closest to one another semantically are the ones found to be grouped together in terms of grammatical expression.

The immediate significance of the analysis presented here is that it has allowed us some insight into the question of what is universal about certain

known grammatical categories, and what aspects of such categories are language-specific. Moreover, these formal categories can now be linked with specific, describable categories of human thought.

In addition, the analysis raises some far broader questions, including the following: What is the role of iconicity and economy in grammatical structure, and to what extent do these and other functional motivations link language with other communication systems? (For some discussion of these motivations in biological systems, see Haiman, 1994). In regard to the degree of elaboration of events, is attunement to this conceptual dimension species-specific; and if so, where does it come from? Is it acquired experientially, and if so, how? Similarly for the apparent limit on the number of categorial distinctions in this conceptual realm: Why is there such a limitation, and in what way does it relate to our general conceptual make-up? Regarding the most general ability supporting these conceptual attunements and constraints, is human beings' capacity for alternate construal of situations shared to any degree by other species? Or is such conceptual flexibility a hallmark of humanity?

I have not answered these broader questions here, but I hope to have succeeded in the more immediate aim of demonstrating the potential for the study of cognitive language typology to shed light on human conceptual capacities and the categorization of experience.

ACKNOWLEDGMENTS

This chapter was prepared in part with the generous support of the Max Planck Institute for Evolutionary Anthropology in Leipzig where I was a visiting research scientist in 2000. I am grateful to Michael Tomasello for sponsoring my visit and to the colleagues at the Institute for a wonderfully stimulating research environment. Thanks also to Michael Barlow for comments on an earlier version of the manuscript and to Maura Velazquez for helpful discussion of Guarani. All errors are my own responsibility.

REFERENCES

- Bowerman, M. (1996). The origins of children's spatial semantic categories: Cognitive vs. linguistic determinants. In J.J. Gumperz & S.C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 145–176). Cambridge: Cambridge University Press.
- Bybee, J., Pagliuca, W., & Perkins, R. (1994). *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.
- Carlin, E. (1993). *The So Language*. (Afrikanistische Monographien 2.) Cologne: Institut für Afrikanistik, Universität Köln.
- Clark, E.V. (1978). Locational, Existential, locative, and possessive constructions. In J.H. Greenberg (Ed.), *Universals of human language: Vol. 4. Syntax* (pp. 85–121). Stanford, CA: Stanford University Press.
- Comrie, B. (1981). *Language universals and linguistic typology*. Chicago: University of Chicago Press.
- Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press.
- Coffin, W. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.
- Faltz, L. (1977). *Referentialization: A study in universal syntax*. Unpublished doctoral dissertation. University of California, Berkeley. (Reprinted 1985, New York: Garland)
- Ferguson, C. (1970). Grammatical categories in data collection. *Working papers in language universals* 4, F1–F15. Stanford, CA: Project on Language Universals, Stanford University.
- Givón, T. (1980). The binding hierarchy and the typology of 'complements. *Studies in Language*, 4(3).
- Givón, T. (1984). *Syntax: A functional-typological introduction* (Vol. 1). Amsterdam: Benjamins.
- Goldberg, A.E. (1998). Patterns of experience in patterns of language. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 203–219). Mahwah, NJ: Lawrence Erlbaum Associates.
- Greenberg, J.H. (1966). *Language universals, with special reference to feature hierarchies*. (Janua Linguarum Series Minor, 59). The Hague, Netherlands: Mouton.
- Greenberg, J.H. (1978). (Ed.), *Universals of human language*. Stanford, CA: Stanford University Press.
- Haiman, J. (1983). Iconic and economic motivation. *Language*, 59, 781–819.
- Haiman, J. (1994). Ritualization. In W. Pagliuca (Ed.), *Perspectives on grammaticalization* (pp. 3–28). Amsterdam: Benjamins.
- Haviland, J. (1979). Guugu Yimidjirr. In R.M.W. Dixon & B.J. Blake (Eds.), *Handbook of Australian languages* (Vol. 1, pp. 24–183). Amsterdam: Benjamins.
- Heath, J. (1978). *Ngandi grammar, texts, and dictionary*. Canberra: Australian Institute of Aboriginal Studies.
- Heine, B. (1997). *Cognitive foundations of grammar*. Oxford: Oxford University Press.
- Hopper, P.J., & Thompson, S. (1980). Transitivity in grammar and discourse. *Language*, 56, 251–99.
- Hopper, P.J., & Traugott, E. (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- Hutchinson, J.P. (1981). *The Kanuri language: A reference grammar*. Madison: University of Wisconsin.
- Kemmer, S. (1988). *The middle voice: A typological and diachronic study*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.
- Kemmer, S. (1993a). *The middle voice*. (Typological Studies in Language, 23) Amsterdam: Benjamins.
- Kemmer, S. (1993b). Marking oppositions in verbal and nominal collectives. *Fatis de Langues*, 2, 85–95.

- Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Langacker, R.W. (1987). *Foundations of cognitive grammar: Vol. 1. Theoretical prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, R.W. (1991). *Foundations of cognitive grammar: Vol. 2. Descriptive application*. Stanford, CA: Stanford University Press.
- Mandler, J.M. (1992). How to build a baby II: Conceptual primitives. *Psychological Review*, 99, 587-604.
- Mandler, J.M. (2000). Perceptual and conceptual processes in infancy. *Journal of Cognition and Development*, 1, 3-36.
- Pawley, A. (1987). Encoding events in Kalam and English: Different logics for reporting experience. In R.S. Tomlin (Ed.), *Coherence and grounding in discourse* (pp. 329-360). Amsterdam: Benjamins.
- Rice, S. (1987). *Towards a cognitive model of transitivity*. Unpublished doctoral dissertation, University of California, San Diego.
- Savage-Rumbaugh, S., Shanker, S.G., & Taylor, T.J. (1998). *Apes, language, and the human mind*. New York: Oxford University Press.
- Slobin, D.I. (1985). Cross-linguistic evidence for the language-acquisition capacity. In D.I. Slobin (Ed.), *The cross-linguistic study of language acquisition: Vol. 2. Theoretical Issues* (pp. 1157-1236). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Slobin, D.I. (2000). Form/function relations: How do children find out what they are? in M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 406-449). Cambridge: Cambridge University Press.
- Stassen, L. (1999). *Intransitive predication*. Oxford: Oxford University Press.
- Sullivan, T.D., & Styles, N. (1988). *Thelma D. Sullivan's compendium of Nahuatl grammar*. (W.R. Miller & K. Dakin, Eds. and Trans.). Salt Lake City: University of Utah Press.
- Talmy, L. (1988). The relation of grammar to cognition. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics* (pp. 166-205). Amsterdam: Benjamins.
- Talmy, L. (2000). *Toward a cognitive semantics: Vol. 1. Concept structuring systems: Vol. 2. Typology and process in concept structuring*. Cambridge, MA: MIT Press.
- Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.
- Tomasello, M. (2000). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Tomasello, M., & Call, J. (1997). *Primate cognition*. New York: Oxford University Press.
- Underhill, R. (1976). *Turkish grammar*. Cambridge, MA: MIT Press.
- Velazquez, M. (1996). *The grammar of possession: Italiendibility, incorporation, and possessor ascension in Guarani*. (Studies in Language Companion Series, 33). Amsterdam: Benjamins.
- Wierzbicka, A. (1992). *Semantics, culture, and cognition: Universal human concepts in culture-specific configurations*. New York: Oxford University Press.

Chapter 4

Social Interaction and Grammar

Cecilia E. Ford

University of Wisconsin

Barbara A. Fox

University of Colorado

Sandra A. Thompson

University of California, Santa Barbara

In this chapter we share some of what we find valuable about the study of grammar as sets of practices adapted to social interaction. Clearly, we are not able to cover all the fascinating and fruitful research that has appeared in this area in recent years, but we hope, through several examples from our own work and through references to other research, to spark further interest in the reader. To begin with, let us consider how we conceive of grammar and then how that relates to the centrality of social interaction as a major habitat to which grammar is adapted.

In interactional settings, we can see grammar "at work." By studying people talking, we can gain a deeper appreciation of what grammar must be understood to be. Three major contributions to our understanding of grammar have arisen from this focus on grammar at work.

The first of these is, in our opinion, one of the most significant contributions to recent linguistic scholarship, a view of linguistic structure itself as rooted in, and shaped by, everyday language use (Bybee, 1995, 1998, 2001, in press; Hopper, 1987; Langacker, 1987). This process of "grammaticalization" is an ongoing one. Thus grammar cannot be a fixed property of human brains, but is emergent, constantly undergoing revision as it is deployed and redesigned in everyday talk.

The second is a recognition that if linguistics is to include an accounting for language in everyday use, then its perspective on the nature of