Abstract

Cognitive Grammar presupposes an inherent and intimate relation between linguistic structures and discourse. Linguistic units are abstracted from usage events, retaining as part of their value any recurring facet of the interactive and discourse context. Linguistic structures thus incorporate discourse expectations and are interpretable as instructions to modify the current discourse state. There are multiple channels of conceptualization and vocalization, including the symbolization of attentional framing by intonation groups. An expression is produced and understood with respect to a presupposed discourse context, which shapes and supports its interpretation. Particular contextual applications of linguistic units become entrenched and conventionalized as new, augmented units. As discourse proceeds, conceptual structures are progressively built and modified in accordance with the semantic poles of the expressions employed. While initially manifesting the specific conceptual structuring imposed by these expressions, the structures assembled undergo consolidation to reflect the intrinsic conceptual organization of the situations described. This conceptual organization has to be distinguished from grammatical constituency, which is flexible, variable, and in no small measure determined by discourse considerations.

Keywords: discourse; Cognitive Grammar; intonation unit; grounding; structure building; constituency.

1. Introduction

The grounding of language in discourse and social interaction is a central if not a defining notion within the functionalist tradition. This is no less true for cognitive linguistics than for other strains of functionalism where the concern is often more visible. My objective here is to elucidate the
connection between the expressions in a discourse, on the one hand, and their structures as described in Cognitive Grammar, on the other. The goal is necessarily modest, since I cannot claim any serious expertise in regard to discourse, nor any extensive familiarity with the vast literature in this area. I will not be presenting any actual discourse data or any detailed analysis. My concern is rather to articulate how Cognitive Grammar and discourse might be brought together, as a matter of principle. I would be very pleased if this were a catalyst for scholars with the proper expertise to investigate their relationship more systematically.

2. Abstraction from usage events

Most fundamentally, Cognitive Grammar makes contact with discourse through the basic claim that all linguistic units are abstracted from usage events, i.e., actual instances of language use. Each such event consists of a comprehensive conceptualization, comprising an expression’s full contextual understanding, paired with an elaborate vocalization, in all its phonetic detail.

The conceptualization inherent in a usage event includes the interlocutors’ apprehension of their interactive circumstances and the very discourse they are engaged in. It thus incorporates their apprehension of the ground (G) and the current discourse space (CDS). The ground consists of the speech event, the speaker (S) and hearer (H), their interaction, and the immediate circumstances (notably the time and place of speech). The CDS is defined as the mental space comprising those elements and relations construed as being shared by the speaker and hearer as a basis for communication at a given moment in the flow of discourse. The ground and the CDS are among the cognitive domains capable of being evoked as the conceptual base for the meanings of linguistic elements.

The various facets of a usage event are spelled out more fully in Figure 1. A usage event is an action carried out by the speaker and hearer. The speaker (S) acts in an initiative capacity, the hearer (H) being responsive; but whether their role is active or reactive, each has to deal with both a conceptualization and a vocalization, the two basic “poles” of an utterance. The speaker’s and hearer’s action involves the directing and focusing of attention (--->). In successful communication, they manage to coordinate this action and focus attention on the same conceived entity. Of course, we can only conceive of so much at any one time. We have a limited visual field, taking in only so much of the world at any given instant. Analogously, we have a limited “conceptual field”, delimiting how much we can conceptualize or hold in mind at any given instant.
Metaphorically, it is as if we are “looking at” the world through a window, or viewing frame. The immediate scope of our conception at any one moment is limited to what appears in this frame, and the focus of attention—what an expression profiles (i.e., designates)—is included in that scope.

A usage event takes place in an immediate context of speech, interpreted broadly as including the physical, mental, social, and cultural circumstances. The ground is at the center of the context of speech, and one element of this context is the very fact that the speaker and hearer are engaged in a coordinated viewing of some facet of the world. In this sense the “window” they are looking through is part of the speech context. Crucially, however, what appears in this window—the content of their coordinated conception—can be anything at all. We are able to conceive and talk about situations in any world (real or imagined), at any time, in any place. The viewing frame can be directed anywhere. The circumstance in which it is directed at the context of speed itself (e.g., This cup is cracked) is just a special (though privileged) case.

Besides the context of speech, the CDS includes a body of knowledge presumed to be shared and reasonably accessible. It also includes the speaker’s and hearer’s apprehension of the ongoing discourse itself: a series of previous usage events, as well as subsequent events that might be anticipated. Any facet of this can be drawn upon or alluded to in the current utterance.

Usage events are bipolar, comprising both conceptualizing and vocalizing activity. Moreover, as shown in Figure 2, at each pole we can reasonably posit multiple channels, which have a certain amount of independence but are nonetheless coordinated in complex ways. Under information
structure, I include such factors as emphasis, discourse topic, and status of information as given versus new. Speech management pertains to such matters as turn taking and holding or relinquishing the floor. The various conceptualization and vocalization channels are shown as occurring within the viewing frame. This is, however, a matter of degree if we understand the viewing frame as delimiting the general locus of attention (the “onstage” region, so to speak). At the conceptualization pole, the objective situation is generally more substantive, more concrete, and more at the center of our attention than the other channels. At the vocalization pole, segmental content is comparably salient.

I have so far talked only about usage events, i.e., actual utterances in an actual discourse context. What about a language and the individual linguistic units of which it consists? To think about linguistic units in isolation from usage events is at best an analytical convenience, and at worst a serious distortion. All linguistic units are abstracted from usage events. Conversely, they are applied to and manifested in usage events. They are best thought of as schematized patterns of action that are immanent in usage events rather than distinct entities. A unit emerges by the reinforcing and progressive “entrenchment” of a commonality (apparent at some level of abstraction) that recurs in a sufficient number of events. A linguistic unit is thus a multifaceted cognitive routine which can be activated and carried out when occasion arises (like the ability to shoot a free throw or to sign one’s name).

What kinds of elements are capable of being abstracted from usage events and conventionalized as linguistic units? Cognitive Grammar takes the straightforward position that any aspect of a usage event, or even a sequence of usage events in a discourse, is capable of emerging as a linguistic unit, should it be a recurrent commonality. A unit can thus incorporate specifications concerning any portion(s) of the overall

\[
\begin{array}{|c|c|}
\hline
\text{Speech Management} & \text{Conceptualization Channels} \\
\hline
\text{Information Structure} & \text{Vocalization Channels} \\
\hline
\text{Objective Situation} & \text{} \\
\hline
\text{Segmental Content} & \text{} \\
\hline
\text{Intonation} & \text{} \\
\hline
\text{Gesture} & \text{} \\
\hline
\end{array}
\]

Viewing Frame

Figure 2.
configuration sketched in Figure 1, including any of the channels detailed in Figure 2. Precisely the same cognitive process—the progressive entrenchment of recurrent commonalities—is therefore held responsible for the emergence of linguistic units of diverse sorts, including those pertaining to pragmatics and discourse. To the extent that such units are conventional in a speech community, excluding them from a “language” or “linguistic system” is arbitrary.

Maximally, therefore, a linguistic unit can correspond to an entire series of usage events constituting a discourse sequence. Figure 3 shows a recurring discourse configuration coalescing as an established unit, thereby being incorporated as part of the linguistic system (i.e., the inventory of schematized patterns of actions controlled by speakers). Since units always originate in this way, in principle they contain all the sectors depicted. However, most units are highly schematic with respect to many or most of these sectors, so that for all practical purposes we can ignore
them. Typically, only certain sectors are characterized with any degree of specificity.

Let us consider a few examples. Figure 4 represents the filler *uh* in English. By means of this device, the speaker can briefly pause while signaling the intent to hold the floor and keep talking. Its relatively specific content is limited to just two channels: the conceptualization channel of speech management, and the vocalization channel of segmental content. Most sectors of the overall structure are highly or fully schematic.

Figure 5 represents the familiar form of the pronoun *you*, in a language like Spanish with a familiar/formal distinction. Conceptually, it profiles the hearer as the focused element in the objective situation. Hence the viewing frame is directed at the ground itself, i.e., a facet of the ground appears onstage as the focus of attention. The hearer’s dual role is indicated by a dotted correspondence line connecting the manifestations of H in two different sectors. In regard to vocalization, this unit specifies the sequence *tu* as its segmental content. Beyond this, as part

---

Figure 4.

---

Figure 5.
of the speaker/hearer interaction in the ground, the unit conveys a relationship of familiarity or solidarity (represented diagrammatically by a double line connecting S and H). All other channels and sectors are fully schematic.

Consider next the sentence-initial use of *therefore*, as in (1).

(1) He’s very rich and powerful. *Therefore* I like him.

As shown in Figure 6, *therefore* introduces and profiles a relationship in which one proposition, labeled P₂, follows from another, P₁. Moreover, P₁ is identified as the proposition expressed in the previous usage event, just earlier in the flow of discourse.

For a last example, consider a strong demonstrative use accompanied by a pointing gesture as in (2).

(2) I want that [°] one.

Besides segmental content, the vocalization (broadly construed, as including all channels of symbolization) includes a pointing gesture on the part of the speaker. Observe, however, that the pointing is not merely a facet of the vocalization, but also part of the expression’s conceptual content—an essential part of its conceptual meaning is that the speaker is pointing and that the intended referent is the entity the speaker is pointing to, typically selected from other contextually available alternatives (given as circles in Figure 7). Hence the speaker has a role in three sectors: as part of the ground, and as the individual pointing in both the gesture channel and the objective situation. Likewise, the entity pointed to functions as such in two sectors.⁵

What about a typical lexical item, e.g., *tree*? I have often described it as a *symbolic unit*, as shown in Figure 8a. Phonologically it comprises a particular segment sequence. Semantically it profiles a thing which has

![Figure 6.](image-url)
various properties abbreviated as T. Yet a fuller description would show these as merely the nonschematic facets of a much more elaborate overall structure, as shown in Figure 8b. Minimally, it shares with other kinds of units the very notion that the expression is being employed by the speaker and hearer, in the context of speech, as part of a discourse. Because this much is common across usage events, the abstracted value of every
unit includes the full schematic framing depicted in these diagrams, even if only a few sectors are specified in any detail.

3. Discourse expectations

Any aspect of a single usage event or a sequence of them in a discourse is capable of being abstracted as a conventional linguistic unit. Thus many units make detailed specifications in regard to multiple events within a sequence. Such units embody established discourse patterns and dependencies. When associated with particular elements, e.g., *therefore* in Figure 6, they amount to discourse *expectations* engendered by those elements. We can find regularities in discourse sequences of any size, viewed at any level of organization. What level we examine determines the scope of the patterns and dependencies described, and what counts as a single usage event for this purpose. It is equally a matter of discourse expectation that a preamble will be followed by a main body of text, that a conditional protasis will be followed by an apodosis, and that a prefix will be followed by a stem. A unified approach is thus adopted for lexical items, grammatical constructions, and discourse patterns of any size.

Following a suggestion by Harder (1996), we might think of linguistic structures (of whatever size) as *instructions* to modify the current discourse space in particular ways. Each instruction involves the focusing of attention within a viewing frame. A discourse comprises a succession of frames each representing the scene being “viewed” and acted on by the speaker and hearer at a given instant. A sequence of frames in a discourse, each corresponding to one updating of the CDS, can be represented as in Figure 9. Heavy lines enclose the frame currently being acted on, which might be called the *focus frame*. I will also call this the *zero frame*, preceded by a *minus frame* and followed by a *plus frame*.

Reference to a sequence of frames is incorporated as part of linguistic units. A particular linguistic element can be *retrospective*, in the sense of making a specification concerning the prior discourse, and/or *prospective*, by virtue of evoking the subsequent discourse. Illustrating a retrospective element, making a specification in its minus frame, is *therefore*, as

![Figure 9.](image)
seen in Figure 6. Illustrating a prospective element, making a specification in its plus frame, would be a topic marker (e.g., Japanese wa), diagrammed in Figure 10. A topic marker refers schematically to the thing profiled by the noun phrase with which it combines; this nominal referent is represented as a circle in the focus frame. Such a marker is prospective by virtue of signaling that the profiled entity will function as a conceptual *reference point* for purposes of interpreting a subsequent proposition, as shown in the plus frame. But it is also retrospective, in the sense that the topic needs to be an entity already accessible in the prior discourse. Thus the profile is shown as corresponding to a thing present in the minus frame.

The size and nature of these discourse frames depends on what level and dimension of linguistic organization we examine. There are discourse dependencies of one sort or another at every level, from the most local to the most global, and to one extent or another we attend to them all, more or less concurrently, when engaging in coherent discourse. Moreover, the same structural element can carry discourse expectations at multiple levels. Figure 6, for instance, shows *therefore* as inducing the expectation of a proposition that would normally be expressed in the previous sentence, as in (1). But at a lower level of organization, *therefore* also carries the expectation of a following clause, which it serves to introduce. Each of these expectations is a schematized remnant of the discourse contexts—at different levels of structure—from which the element in question is abstracted by the reinforcing of recurring commonalities. Each constitutes a structural frame (or constructional schema) which the element participates in and evokes as part of its conventional linguistic value.

Consider the constructional schema for *if* … *then* sentences, sketched in Figure 11. *If* is taken as designating the selection of one proposition, P₁, on a hypothetical basis, from a range of possible alternatives—i.e., it evokes a hypothetical mental space. It further induces the expectation that another proposition (P₂) will be valid in that space. *Then* designates the relationship wherein one proposition, P₂, follows from another, P₁ (cf. Dancygier and Sweetser 1997). At the lowest level of constituency, *if* combines with a finite clause, which expresses a proposition and profiles

![Diagram](image_url)
a grounded process abbreviated as Pr. Likewise, then combines with a finite clause profiling the process given as Py. In each case the clause imposes its profile at the composite structure level, which is why we refer to these intermediate-level structures as an “if-clause” and a “then-clause” (not as complex connectives). I will assume that when the protasis and apodosis combine to form the highest level composite expression, it is the apodosis whose profile prevails.

We can look at Figure 11 in various ways, all equally valid. It is a constructional schema describing a grammatical pattern, which is also a local discourse pattern. It can further be seen as part of the lexical characterization of if, as well as then. Since lexical items occur in particular structural frames, and are acquired by abstraction through these occurrences, schematized representations of a set of such frames are retained as part of a lexical item’s conventional linguistic value. The description of if, for instance, includes its role in the constructional schema specifying its combination with a finite clause, as well as the higher-level constructional schema specifying how this finite clause combines with the conditional apodosis (i.e., the entire structure in Figure 11). The prospective or retrospective nature of a lexical item resides in its role in these constructional schemas. Diagrams like Figures 9 and 10 (with plus, minus, and focus frames) can be thought of as abbreviations for such schemas.
4. Attentional framing

Classically recognized constructions, like the one in Figure 11, pertain primarily to the objective situation and segmental content. Clearly, though, a comprehensive linguistic description has to accommodate structures and relationships in all relevant channels. We have already seen examples involving gesture (Figure 7), speech management (Figure 4), and information structure (Figure 10). With respect to the latter, let us examine more extensively a dimension of organization involving the management of attention in the flow of discourse, especially at a consciously accessible level. The frames in question are thus “windows of attention”—single attentional “gestures”—typically marked by intonation.

These frames can, I believe, be identified with Chafe’s notion of an intonation unit:

Spoken language lends itself to segmentation into intonation units. Such units are identifiable on the basis of a variety of criteria, among which are pauses or breaks in timing, acceleration and deceleration, changes in overall pitch level, terminal pitch contours, and changes in voice quality. Intonation units are hypothesized to be the linguistic expression of information that is, at first, active in the consciousness of the speaker and then, by the utterance of the intonation unit, in the consciousness of the listener... [S]ubstantive units are fairly strongly constrained to a modal length of four words in English, a fact that suggests a cognitive constraint on how much information can be fully active in the mind at one time... [T]he majority of substantive intonation units have the form of single clauses, though many others are parts of clauses. (Chafe 1994: 69)

The frames I will be dealing with correspond to actual or likely substantive intonation groups, in Chafe’s terms. They are expressions of modest size, typically comprising one to several words; these are phonologically cohesive, often set off by slight pauses; conceptually they consist of information fully active in the mind at one time; grammatically they tend to coincide with single clauses.

This dimension of linguistic organization involves the presentation of information in coherent “packets” of digestible size. Chafe indicates that the majority of substantive intonation units correspond to single clauses. I believe this represents a very natural mapping, especially in the case of finite clauses, which incorporate grounding and thus indicate how the speaker and hearer view the profiled process in relation to their own circumstances. It makes sense that such an assessment should tend to coincide with a single window of attention. Still, we often cram more than one clause into a single frame, or stretch out a single clause into more than one frame. By doing so we can regulate the amount of effortful
material presented in a single window of attention. Chafe noted a “one new concept at a time” constraint, “the finding that only one concept can be changed from the inactive to the active state” during any one intonation group (Chafe 1987: 31). Also related is the tendency noted by Du Bois (1987: 819) and others for at most one new participant, lexically expressed, to be introduced in a single clause.

I will refer to this dimension of organization as *attentional framing.* The discourse frames in question are bipolar: at the phonological pole each consists of some kind of intonational grouping; and at the semantic pole, each comprises a body of active information all of which is simultaneously available. They are also *symbolic,* in that the very act of imposing a particular intonational grouping effects and symbolizes the act of imposing the corresponding conceptual grouping.

As sketched in Figure 12, the semantic pole of an attentional frame belongs to the channel of information structure, and its phonological pole, to intonation. An attentional frame comprises no specific conceptual or segmental content. Its conceptual value resides in the very act of making a single attentional gesture—imposing a single window of attention for the simultaneous viewing of conceptual content. Likewise, its phonological value resides in the very act of imposing an intonational grouping. Despite its abstract nature, an attentional frame is symbolic to the same extent as a lexical item like *tree.* The difference is merely

![Figure 12.](image-url)
that their primary semantic and phonological specifications pertain to
different channels.

Structurally, expressions consist of complex assemblies of symbolic
elements, in large measure arranged in hierarchies of component and
composite structures. In terms of attention management, they comprise
(at the very least) a linear sequence of successive attentional frames. In
putting together an actual expression, in all its organizational dimen-
sions, a mapping has to be established between elements in the structural
and attentional domains. We can speak of mappings between separate
“tiers” (as in phonology); here I adopt the metaphor of “planes”. In
any case, each “tier” or “plane” is bipolar, incorporating at least one
“channel” (to use yet another metaphor) at the semantic pole and one at
the phonological pole.

An instance of this mapping is shown abstractly in Figure 13(a), where
capital letters are used for conceptual structures at the semantic pole, and
small letters for phonological structures. Dotted correspondence lines
indicate mappings between successive attentional elements, whose
arrangement is basically linear, and particular structural elements,
whose arrangement is both linear and hierarchical. The result of merging
the structural and attentional structures in the way specified is a kind of
composite representation, shown in Figure 13(b), where the attentional
framing is superimposed on structural elements, as indicated by the
heavily outlined boxes.

Of course, there need be no single, unique mapping for a given
expression. In Figure 14 (where the semantic and phonological poles are
collapsed, and hierarchy is represented by nested boxes rather than a tree
diagram) four alternate attentional framings are shown as being imposed
on the same structure. A primary means of imposing alternate framings
is by equating individual frames with elements at different levels in the
structural hierarchy. For the same expression, this will often result in
different numbers of frames. As a special case, it is possible for a single
frame to subsume the entire expression, as in diagram (c).

Consider two alternate framings for an if … then sentence, as in (3):

(3) a. If she said it, then it’s true.
   b. If she said it then it’s true.

These are indicated by the heavily outlined boxes in Figure 15 (which
corresponds to the upper portion of Figure 11). The contrast is a matter
of whether the overall expression is parsed into two attentional units,
representing the intermediate-level composite structures (the if-clause
and then-clause individually), or whether the entire expression is
squeezed into a single attentional frame.
The basic meaning is the same with either framing. In either case, the full expression has the same overall conceptual content, and conditionally asserts $P_y$, just as it incorporates the same overall segmental content. Still, the alternate framings do impose—indeed, they constitute—a subtle semantic contrast. Dwelling on each clause individually, in a separate attentional gesture, enhances its cognitive salience and that of the elements it contains, if only by according them more processing time and thus
a fuller realization. On the other hand, incorporating both clauses in a single attentional gesture results in their phonological and conceptual “compression”—more has to be squeezed into a single, limited span of processing time, resulting in a somewhat less articulated realization. It is quite analogous to the difference between typing two letters, say K and L, in two separate gestures, versus typing them together in one coordinated motion. Both get the job done, in the sense that both letters get typed. Yet the two implementations are qualitatively quite distinct, both experientially and at the level of motor activity.

I would strongly deny that the meaning contributed by attentional framing lies outside the proper scope of linguistic semantics. Linguistic meaning is largely a matter of construal, and dimensions of construal reasonably described as matters of prominence are critical to both semantics and grammar. The conceptual compression and reduced salience of individual elements resulting from their being incorporated in a single attentional frame are no more mysterious and no less real than their phonological analogs, which everyone would acknowledge. This is comparable to the reduced salience of components that occurs when a complex expression undergoes a progressive loss of analyzability, so that we no longer dwell on the components individually (although they are still there). For instance, the notions ‘ice’ and ‘cube’ are much less salient in the familiar compound ice cube than they were when the expression was first formed, or than they are currently in a novel expression like small cube made of ice.

Moreover, particular attentional framings become entrenched and conventionalized as part of a linguistic system. Both framings in example (3) are familiar, fully conventional ways of implementing the if…then construction. The configurations shown in Figure 15, including their attentional framings, represent complex constructional schemas learned and used by speakers of English. If we posit a more schematic
constructional schema, neutral in regard to framing, it can be seen as a further abstraction from the attentionally framed expressions a speaker actually encounters in discourse.

I am suggesting, then, that it is common for a particular kind of structure to occur with multiple attentional framings, each of which—in conjunction with the structure—may constitute an entrenched, conventional linguistic pattern. One particular framing may be unmarked, or
prototypical, but an array of alternatives may all be familiar and conventionally sanctioned. This family of alternate framings is part of the full characterization of the structure. I think the situation is quite analogous to that of a lexical item and the set of structural frames it conventionally occurs in: schematic reference to those frames is part of its overall characterization, and a frame-neutral representation arises (if at all) by abstracting away from their differences. Indeed, pushing things one step further, attentional framings conventionally imposed on those structural frames can also be considered part of the characterization of a lexical item. I would argue, for instance, that the full configurations in Figure 15 are conventionally associated with if and thus belong to its linguistic characterization, in the fullest sense. Whether we view these configurations as pertaining to lexicon, grammar, or discourse is simply a matter of the broader context (the more inclusive network of related structures) in which we choose to examine them.

Even within a single clause, particular specifications about attentional framing are often incorporated as conventional facets of grammatical constructions. For instance, English has two conventional framings for nominal apposition. The contrast between two successive fixations of attention, and a single fixation in which all the information is simultaneously active, accounts neatly for the well-known distinction in example (4a). The single-frame construction does not merely offer two independent characterizations of the same individual, but offers a single, complex characterization in which the first specification is of a more personalized nature (reflected in the possessive) and the second more objective. The two-frame construction is thus more easily reversible and allows a much broader range of nominals, as seen in the other examples.

(4) a. (i) my son, the doctor (ii) my son the doctor
b. (i) her friend, Jane Doe (ii) her friend Jane Doe
c. (i) Jane Doe, her friend (ii) *Jane Doe her friend
d. (i) our accountant, Kim Kelly (ii) *our accountant Kim Kelly

Consider next the complex locative construction in example (5), where the first locative element specifies a general location, and the second a more specific location nested within it. Particular locatives— notably upstairs, downstairs, inside, and outside—occur in a constructional variant where they belong to the same intonation group as the one that follows, as in (5a) and (5b):

(5) a. It was upstairs, in the attic.
b. It was upstairs in the attic.
c. *It was up, in the attic.

d. It was up in the attic.

In (5a) *upstairs and in the attic are treated as two separate locations, which happen to be nested, whereas (5b) presents them as two facets of a single, complex locational specification. The shorter forms *up, *down, *in, and *out are limited to the latter configuration. Like “phrasal prefixes”, they are bound to the following locative.

It is well known that an adverbial clause can either occur as a separate intonation group or join the main clause in a single group:

(6) a. He married her, because she’s intelligent.
    b. He married her because she’s intelligent.
    c. He didn’t marry her, because she’s intelligent.
    d. He didn’t marry her because she’s intelligent.

Though (6a) and (6b) appear to be semantically equivalent, they differ in that the two-frame expression presents the main clause event independently, with a subsequent explanation for its reason, whereas the single-frame expression presents them as part of a single judgment. In the most likely interpretation, (6b) presupposes that he married her and asserts that her intelligence was the reason. This semantic contrast becomes more obvious when the main clause is negated. In (6c), the focus of negation is the main clause event itself, since the reason is presented as a separate statement. Strikingly, however, the favored interpretation of (6d) is that he did in fact marry her, even though the main clause bears the negative marking. Since the entire sentence expresses a single complex judgment, in which the main clause event is presupposed, the force of negation falls on the assertion that her intelligence was the reason.

I might also mention the intonational contrast between restrictive and nonrestrictive relative clauses. A nonrestrictive relative is typically manifested as a separate intonation group:

(7) a. That car which I just repaired is obviously very expensive.
    b. That car, which I just repaired, is obviously very expensive.

Its semantic property of not restricting the head noun’s possible referents is quite consonant with its status as a separate attentional gesture.

So attentional framing cannot be dismissed as merely a performance-induced artifact of spoken discourse. We learn and use conventional framing patterns, as well as conventional ways of mapping them onto structures of different sizes. Particular framings are incorporated as integral facets of grammatical constructions. There is, to be sure, a great deal of flexibility in how we do the mappings, but even in its most
spontaneous, ungrammaticized manifestations the framing is partially shaped by convention, a matter of constrained freedom rather than unbridled license.

While attentional framing tends to correlate with constituency, it cannot be reduced to it. For instance, the contrast in (3) cannot be described just by positing two different phrase trees. With either attentional framing, we have to say that the *if*-clause and the *then*-clause are separate grammatical constituents that combine to form a single, larger constituent (in this case the sentence as a whole). The contrast does not reside in constituency per se, but is rather a matter of which *level* in the constituency hierarchy is selected for simultaneous examination in a single attentional and intonational gesture. Nor do intonation groups regularly correspond to constituents traditionally recognized on other grounds. For instance, the division of (5a) into the major constituents *it was upstairs* and *in the attic* would not conform to any standard practice. Intonational groupings of nonconstituents, or ones that cross-cut seemingly evident constituency hierarchies, are readily found. Chafe recorded (8a) in an actual conversation. Though (8b) is a concocted example, its phrasing seems quite natural to me.

(8)  

a. Have the ...áanimals, ever attacked anyone in a car? (Chafe 1994: 61)  
b. I admire people who even, when they are very busy, are willing to help out.

The relationship I see between attentional framing and grammatical constituency is therefore not a simple one. I would posit framing patterns that are independent of any particular constituency, and can thus be flexibly imposed on structures of different kinds and different sizes. I would also posit constituency hierarchies that are independent of any particular framing (e.g., the *if*...*then* construction supporting the alternate framings in Figure 15). At the same time, particular mappings between the two are conventionalized and incorporated as integral facets of grammatical constructions, at various levels of organization and specified at different levels of abstraction. Whether the mappings are novel or conventionalized, there is a natural tendency for attentional frames to coincide with grammatical constituents, but in either case it is only a tendency, not an inviolable principle.

I have argued elsewhere (1995, 1997) that constituent structure, as traditionally understood, is neither rigidly fixed nor essential to grammar. To the extent that it emerges, it can be seen as a special case of the broader, more basic phenomena of conceptual grouping, phonological grouping, and symbolization. A classical constituent emerges when a particular kind
of conceptual grouping, based on an overlap involving salient elements, happens to be symbolized by a particular kind of phonological grouping, based on linear contiguity. But there are many kinds of conceptual and phonological groupings, which are linguistically important irrespective of whether they participate in symbolic relationships of the classical or any other variety.

I therefore have a much broader than usual view of what can emerge as either a classical or a nonclassical constituent. I might well be willing to recognize *it was upstairs and in the attic* as the major constituents in (5a), or the two intonation groups as the ones in (8a). At the same time, if the notion of constituent is interpreted quite broadly, as any symbolic pairing between a conceptual and a phonological grouping, then by its very nature attentional framing imposes a kind of constituency, whether it coaligns with classical constituency or cross-cuts it. For instance, in (8b), I would not consider *I admire people who even* to be a classical constituent; that is, I do not think it is assembled first as a self-contained structure based on valence links, then combined with the remaining portions of the sentence. It is however a constituent, in the broad sense, because the attentional framing itself imposes a pairing in which a phonological grouping—the intonation group—symbolizes a conceptual grouping consisting of the elements presented in a single window of attention.

5. Application to the CDS

Let us come back to the notion of viewing linguistic structures as instructions for manipulating the current discourse space. As discourse unfolds, the expressions successively encountered serve to update the CDS. When attentional frames correspond to clauses, major updating of the CDS proceeds on a clause-by-clause basis. Still, I presume that updating takes place on numerous time scales, corresponding to different levels of structural and conceptual organization.

Whatever the level of organization, an instruction to update the CDS involves at least two discourse frames: a minus frame corresponding to the “input” CDS, i.e., the one to be updated; and the zero or focus frame, the one being created by the updating process. Even in degenerate cases, where the minus frame makes no significant specifications, it has at least an implicit schematic presence as the structure being operated on. For example, the noun *triangle* is partially characterized as shown in Figure 16. Without imposing any particular conditions on the prior discourse state, it constitutes an instruction to focus attention on a thing of a certain sort, i.e., to profile it.
It is a truism that linguistic expressions are always produced and understood relative to some presupposed discourse context. Unlike many truisms, I suspect that this one may actually be true. As a lexical item, triangle specifies the shape of its profile only schematically. When actually used in context, however, it may well be understood in a very specific way. Suppose a particular triangle is already salient in the CDS (perhaps it is visible on the blackboard in a geometry class). There will then be a strong tendency, in using the expression triangle, for the entity put in focus to be identified with the one already present in the CDS.

This is sketched in Figure 17, representing the application of an expression to the CDS as part of an actual usage event. The expression, triangle, is shown at the bottom. Its minus and zero frames are aligned with phases of the CDS, and the instructions it embodies are carried out in this context. In the case at hand, the triangle put in focus is identified with the one already present in the CDS, as indicated by the diagonal correspondence line. (If none were available, of course, the expression would serve to introduce it.)

The structure on the upper right, the focus frame in the evolving CDS, represents the way in which triangle is actually understood in the context indicated—the expression’s actual contextual understanding in a possible usage event. This contextual understanding is a function of both the lexeme’s abstracted meaning and the CDS to which it is applied; it is probably never precisely the same on any two occasions, since the CDS is never precisely the same. Naturally this is also true of complex expressions, fixed or novel. It is a major reason why language exhibits only partial compositionality (Langacker 1998).

6. Augmentation

Like any other facet of usage events, a particular way of applying a linguistic unit to the CDS can itself—should it recur—be reinforced, entrenched, and conventionalized. By definition, this results in a new, augmented linguistic unit. The basic unit is thus a component of the
augmented unit, but may also remain accessible for independent use. This fundamental mechanism of language change is represented abstractly in Figure 18.

Because it resides in a particular way of applying an expression to the CDS, augmentation does not per se result in any change in form. While the augmented structure looks the same as the basic unit, it incorporates additional specifications that are “pragmatic” in origin, but are by definition “linguistic” once the application is entrenched and conventionalized. I believe a great deal of linguistic structure is “invisible” because it resides in conventionalized ways of applying more directly observable structures to the CDS. Invisibility does not make it any less real or important, however. As conventional symbolic structures that are an integral part of one’s ability to speak a language, these augmented units can be excluded from the linguistic system only arbitrarily (Langacker 1998).

An example is the conventional pattern of English whereby a declarative finite clause, without special marking, can be used to make an assertion. A declarative clause per se does not include assertion (or any other speech act) as part of its inherent value. Observe that the proposition *Harold has finished his thesis* is not asserted in any of the expressions in (9):

(9) a. Conceivably, Harold has finished his thesis.
   b. It’s not the case that Harold has finished his thesis.
   c. Jenny suspects that Harold has finished his thesis.
   d. Harold has finished his thesis. And I was just elected pope.

It is only when a finite clause is embedded in a larger interactive frame—involving the speaker’s intent to portray the proposition as true...
and offer it for the hearer’s assessment—that producing it constitutes an assertion. If the schematic description of a declarative finite clause is seen as a basic unit, we can posit an augmented unit that specifies in schematic terms the use of such a clause in those interactive circumstances.

Patterns of pragmatic inference can also become entrenched and conventionalized as linguistic units. Consider example (9d). In this familiar pattern, the speaker pretends to make a plausible assertion, but follows it with another apparent assertion which is so obviously false that the hearer realizes something other than real assertion is going on, drawing the conclusion that the first sentence must actually be false as well. Whether or not this conclusion follows from general pragmatic principles, it constitutes an entrenched pattern that speakers learn and employ as a matter of established convention. The discourse schema representing this pattern incorporates the augmented unit for assertion, just described, and embeds it in a more encompassing interactive frame, which specifies that the proposition apparently being asserted is actually being denied. This embedding produces a higher-order augmented unit, as shown in Figure 19. Being a conventionally sanctioned facet of our ability to speak and understand English, this higher-order unit is part of the language (unless excluded by arbitrary definition).

7. Augmentation and grounding

Among the specifications commonly effected tacitly, by means of augmented units, is *grounding*. By grounding I mean the grammaticized specification of the relation between the ground and the profile of a noun phrase or a finite clause in regard to certain fundamental “epistemic” domains, such as time, reality, and coordinated mental reference. A noun or a verb by itself merely specifies a *type* of thing or process. A full noun phrase or a finite clause profiles a *grounded instance* of a thing or process type (Langacker 1991). Clausal grounding elements in English are tense and the modals. The nominal grounding elements include the articles, demonstratives, and certain quantifiers.
A common noun, e.g., triangle in Figure 16, is merely an instruction to direct attention to a certain type of thing; it does not itself single out any particular instance of that type, nor does it invoke the ground in any special way. On the other hand, a grounding element provides a means for the speaker and hearer to establish coordinated mental reference with an instance of some type. Grounding is a primary means of specifying how the content of an expression is to be integrated in the CDS and connected to the conceptual structure already in place. The reference it establishes is reference in the context of the discourse.

Figure 20 introduces abbreviatory notations for certain kinds of nominal elements. As indicated, a noun merely focuses attention on a thing of a certain type (t), without singling out any instance. A grounding predication, such as the definite or indefinite article, does invoke the ground, notably the speaker and hearer. In diagrams (b) and (c), solid lines connect the speaker and hearer with a thing inside the viewing frame. These lines indicate that the speaker and hearer achieve mental contact with a particular thing instance, i.e., they single it out for individual conscious awareness.

An indefinite article, a definite article, and a personal pronoun all result in the speaker and hearer directing attention to a particular instance of a thing type (hence the focus frames in Figure 20[b] to [d] are all the same). The difference among them resides in how this happens. The indefinite article implies that no unique instance of the type in question is yet accessible in the CDS. Either an instance has to be newly introduced in the CDS, as in example (10a), or one of multiple instances has to be selected and put in focus, as in (10b).

(10) a. On the desk, he noticed an important-looking document.

b. There were two mice and a rat in the cage. I picked up a mouse and examined it.
The definite article implies that a unique instance of the specified type is indeed accessible in the CDS. The referent is thus in some way already available in the minus frame, as indicated by the correspondence line in Figure 20(c). It is not however required that it was previously mentioned or even salient, or that coordinated mental reference to it had been previously achieved; it need only be achieved by the definite nominal itself. Observe the examples in (11):

(11)  a. There were two mice and a rat in the cage. I picked up the rat.
    b. I’m worried about my car. The engine is making funny noises.
    c. It’s quiet. I think the air-conditioning just went off.

Finally, a personal pronoun implies that its referent is not only accessible in the CDS but has in some manner already been singled out. It must also remain salient enough that it will offer itself as the intended referent given only the schematic content of the pronoun (van Hoek 1995, 1997).

We must next consider how a grounding element combines with a nominal expression to form a full noun phrase. This is shown in Figure 21 for the case of an indefinite article; the component structures are given at the left, the composite structure at the right. Observe that the respective minus and zero frames of the component structures correspond to one another. In other words, the article and the noun are in sync with respect to the flow of discourse (at least when viewed at a global level of organization). Together they yield a composite structure which functions as a single discourse unit, a “referential gesture” serving to introduce and focus attention on a particular individual, subsequently available as an established discourse referent.
Consider next the application of this composite structure to the CDS. Figure 22(a) represents a proper usage of the indefinite nominal *a dog*, which implies that no unique instance of the dog category is yet accessible in the CDS. In the case depicted, the minus frame of the CDS makes no reference at all to a dog, so this condition is satisfied. The nominal constitutes an instruction to focus on an instance of this category, which means in this circumstance that such an instance has to be introduced into the discourse, where it appears as the focused (profiled) element in the zero frame of the CDS. The updated CDS results from carrying out the instructions embodied in the nominal expression in the context of the previously existing CDS (the minus frame).

In Figure 22(b) we see a proper use of a nominal with the definite article. Inherited from the definite article is the specification that a unique instance of the thing type (*dog*) is already available in the CDS, though perhaps not previously singled out for coordinated reference. The minus frame of the CDS contains such an instance, so the nominal can be used felicitously. By carrying out the instructions it embodies, the uniquely available instance of the thing type *dog* is singled out as the focus of attention in the updated CDS (the zero frame).

What about a language that does not have definite and indefinite articles? Many languages make do with an occasional use of demonstratives and the numeral ‘one’, leaving many nominal expressions unspecified for definiteness/indefiniteness. An example is Luiseño (a Uto-Aztecan language of southern California):

(12) *Xosee ‘awaal-i ‘ari-qu$.
José dog-OBJ kick-PAST:DUR
‘José was kicking {the/a} dog.’
What do we say about nominal grounding in cases like this, where there is no visible grounding element?

I think we can reasonably posit augmented units. It is a conventional fact of Luiseño that a nominal expression without any overt grounding element can be used in discourse with either definite or indefinite reference. These conventional usage patterns take the form of schematic augmented units specifying how a noun—the basic unit—can be applied to the CDS in a usage event. Let us suppose that a noun is correctly described as shown in Figure 20(a). This noun schema functions as the basic unit. Two augmented units based on it, both conventionally established in a language like Luiseño, are sketched in Figure 23. In diagram (a), an instance of the specified type (t) is introduced into the focus frame of the CDS as a new discourse referent. In diagram (b), an instance is already available in the minus frame of the CDS, so that instance offers itself as the one singled out and profiled in the focus frame. Of course, we can also posit a higher-level schema that abstracts away from these two augmented units, being neutral as to the previous discourse status of the profiled referent. The point is that the two options are conventionally sanctioned, as general patterns available for any noun.11

Augmented units such as these, specifying conventional patterns for applying structural elements to the ongoing discourse, require no additional theoretical apparatus in Cognitive Grammar. The configurations in Figure 23 are just symbolic assemblies of a particular kind, i.e., configurations of conceptual structures, phonological structures, and symbolic linkages between them; they are special only in the sense that some of these conceptual structures are identified with the CDS. But I have

---

Figure 22.
always maintained that one of the cognitive domains figuring in the characterization of linguistic elements is the conception of the ongoing discourse itself. Moreover, structures like these arise by schematization capturing what is common across a series of usage events, and are immanent in particular events employing them. Nothing new is being posited—I am only pointing out a certain way of exploiting the descriptive potential of the framework.

8. Structure building

We have considered individual examples of how a linguistic expression is applied to the CDS and updates it in some fashion, e.g., Figures 22 and 23. What remains is to extend this notion to connected discourse, where each successive utterance updates the CDS, usually by building on what is already established there. This is obviously a vast subject, on which much work has been done from many perspectives. The most I can do is offer some limited remarks on how particular matters might be approached from the standpoint of Cognitive Grammar. I will start by briefly mentioning a few general principles of structure building, none terribly new or original.

The basic units of structure building are relationships rather than things. This reflects the fact that connected discourse normally comprises a series of clauses (each profiling a relationship). Only in very special circumstances does a coherent discourse consist in a series of nominal expressions.

When structure is added, it is usually anchored to what has already been built, by virtue of conceptual overlap. This is implied by the very notion
of building a coherent structure. Familiar manifestations include the given/new distinction, as well as anaphora and other kinds of retrospective elements. In the diagrams used here, e.g., Figure 23(b), it is reflected in correspondences between facets of linguistic structures and minus frames in the CDS.

Even when covert, the ground is the ultimate anchor for building connected structures. Observe that things and relationships may be directly linked to the actual ground via a grounding predication even when they function as parts of special mental spaces. In (13a), for example, this and is anchor the profiled thing and process to the actual ground despite their role in the mental space representing the report of outside reviewers. In (13b), where a structure is being built in the fictive world of a movie, the ground is still the ultimate anchor, since the movie in question is linked to it via the demonstrative that. (In complex cases, of course, the connection with the actual ground may be very indirect.)

(13) a. The reviewers wrote that this paper is terrible.
    b. In that movie, the Easter Bunny dreams of being kidnapped by aliens.

In a sequence of clauses, a subject has special status as point of attachment to what has already been constructed and point of access to what is currently being constructed. Its usual status as point of attachment to the structure already in place relates to the well-known tendency for a subject to be given rather than new in the discourse. Its description as point of access to what is currently being constructed alludes to Chafe’s characterization of a subject as a starting point to which other information is going to be attached (Chafe 1994). This is consonant with my own characterization of a subject as trajector or primary focal participant, i.e., the initial participant invoked for purposes of arriving at the conceptualization of the profiled relationship (Langacker 2001). To take just one example, sentence (14) is a natural continuation of the Luiseño example in (12):

(14) Xaari-qu$.
    growl-PAST:DUR
    ‘It [the dog] was growling.’

The subject (the dog) can remain implicit because it has already been established in the discourse, providing the point of attachment for the added structure concerning the event of growling. The dog is also the point of access to that event, being identified with its trajector, the starting point for conceptualizing the profiled process of growling.
Certain grammatical positions are much more commonly used than others for the introduction of new discourse participants. Two asymmetries in particular can be noted. First, as Du Bois (1987) and others have observed, a new participant is much more likely to be introduced as an object than as a (transitive) subject. Second, a new participant is seldom introduced as a possessor, but often in the role of possessed.

Starting with possessives, let us first observe the virtual impossibility of introducing a new participant in a possessor role. For instance, (15a) is quite infelicitous as a way of first introducing a lion into the discourse. On the other hand, if a lion has previously been introduced, as in the first sentence in (15b), it can function as possessor in a construction that first introduces the mane as a discourse referent, in the role of possessed.

(15) a. *A lion’s mane was long and tangled. It roared.
    b. We found a sleeping lion. Its mane was long and tangled. We tried to comb it out.

I would explain this asymmetry in terms of the conceptual characterization of possessive constructions that I have argued for elsewhere (1993). Specifically, I have characterized a possessor as a conceptual reference point affording mental access to a range of potential targets (its dominion), the one accessed being the possessed entity. The reference point relationship is inherently asymmetrical—the target is mentally accessed via the reference point, rather than the reverse. It stands to reason, then, that possessor and possessed should strongly coalign with given versus new in the discourse.

More controversial is the claim that a subject can also be characterized conceptually as a reference point. I view a subject and object as the first and second conceptual reference points accessed in building up to the full conception of a profiled relationship. If we just examine the relational participants, the subject is the initial reference point, and an object is a target evoked in the context of conceptualizing a relationship the subject participates in. Since the subject anchors the conception of a relationship encompassing the object, it stands to reason that subject and object should coalign with given versus new in the discourse. The first sentence in (16a) observes this natural alignment, with the subject given and a new participant introduced in object position. The second sentence in (16a) is less natural, because it has the opposite alignment; the continuation in (16b) is preferable, even though it uses a passive.

(16) a. *We saw one of the hostages. A mean-looking thug was guarding the poor girl.
b. We saw one of the hostages. The poor girl was being guarded by a mean-looking thug.

Since I characterize all subjects as starting points or reference points, the question arises why only transitive subjects resist the introduction of a new participant. It is well known that the role of intransitive subject is commonly used for this purpose (Du Bois 1987; cf. Herring 1989). The basic reason, I believe, is that many nontransitive clauses are specifically tailored to facilitate the introduction of a new participant. Often the verb is one of motion, posture, or existence, so that its meaning directly subserves the function of establishing its subject in the scene, as in (17a).

(17) a. A lion strode into the clearing. Its mane was long and tangled.
    b. Into the clearing strode a lion. Its mane was long and tangled.
    c. There was a lion in the clearing. Its mane was long and tangled.

Beyond this, devices like the locative inversion in (17b) allow the initial specification of a location as a kind of conceptual anchor, so that even though the subject is new it can attach to something already introduced. I analyze elements like the “existential” there in (17c) as having a comparable function. It profiles an abstract setting, a “presentational frame” in which a new participant is about to be introduced. Its status as grammatical subject reflects this function, since a presentational frame is very naturally adopted as a conceptual starting point, reference point, or point of mental access.

Let us now take a coherent discourse and consider how it builds a connected conceptual structure, step by step. The discourse in (18) is trivial and obviously concocted. It is only intended to give convenient, concrete illustration—for the simplest sort of case—of how structure building in discourse relates to the structure of linguistic expressions as described in Cognitive Grammar. We can assume, in this simple kind of case, that the primary updating of the CDS takes place on a clause-by-clause basis. Recall that a clause is the typical unit chosen for a single window of attention symbolized by an intonation group.

(18) a. I was sitting in the park yesterday.
    b. A man walked by.
    c. The man saw a dog.
    d. He kicked it.
    e. It yelped.

Working at this level of organization, we can take each clause as providing a single conceptual “chunk” employed for updating purposes.
This chunk can be identified as the clause’s composite semantic structure, comprising (at least) a focus frame containing the processual profile, as well as a minus frame specifying any restrictions it imposes on the CDS. The CDS, at successive stages in this discourse, is represented in the middle row in Figure 24. The structures in the top and bottom rows are the conceptual chunks doing the updating, i.e., the composite semantic structures of the sentences in example (18b) to (18e). Basically, in one updating cycle, the minus frame of the composite expression is checked against the CDS, as already constituted. If the CDS satisfies the requirements of the composite expression, the instructions embodied in the expression are carried out in that context, resulting in a new, updated phase of the CDS. This is then checked against the minus frame of the next expression, leading to a further updating, and so on indefinitely.

We start, at the middle on the left, with the CDS as it stands after the processing of (18a). Not depicted, this initial expression sets the stage by specifying a time and place—in relation to the ground—where the connected discourse structure is going to be built.

The next clause, (18b), profiles a process (walk by) that brings a new participant onto the scene. The indefinite article indicates that a unique instance of the type man is not already present in the CDS, a condition that is satisfied. The clause thus applies to the CDS so as to introduce the participant as trajector of the profiled event. In the updated CDS, a previously unmentioned man is therefore focused upon as trajector in a profiled instance of walking by, situated in the spatial and temporal context already established.

The following clause, (18c), does require of the CDS to which it applies a unique instance of the man category, and the absence of such for the dog category. These conditions are satisfied, so the clause applies to the CDS, producing an updated CDS which profiles an event of seeing involving the man, as trajector, and a newly introduced participant, a dog, as landmark. Observe that the prior event (of the man walking by) does not abruptly disappear from the CDS. But it is no longer in focus, as it is not profiled in the focus frame of the clause currently being processed. Naturally, as one event is put in focus at a given processing phase, events that were focused previously decline in salience and may eventually fade out altogether, no longer being available in the CDS.

The building process continues with (18d). This clause constitutes an instruction to introduce and profile a force-dynamic event involving two participants. Since the trajector and landmark are both expressed by personal pronouns, their referents must already be accessible in the CDS, and have previously been singled out by the speaker and hearer (van Hoek 1995, 1997). The schematic semantic values of the pronouns he and
Figure 24.
are compatible with the characterization of two such referents, the man and the dog, respectively. Making these identifications results in the fourth phase of the CDS, which profiles an act of kicking.

Finally, (18e) would normally be interpreted as referring to the dog’s reaction. The pronoun *it* is sufficient to identify the trajector, since the CDS contains only one salient non-human participant. Although the diagram portrays all the previous events as still being active in the CDS, by now the earlier events have no doubt begun to fade.

Oversimplified, sanitized, and non-naturalistic though this example may be, it does illustrate how a series of utterances can build a coherent, connected conceptual structure by successively updating the CDS. Describing the conceptual structures being progressively built up, and how they relate to the meaning and grammar of individual expressions, is essential to a clear and comprehensive understanding of discourse and how it unfolds. A detailed characterization of the conceptual structures being built and manipulated is as fundamental and indispensable to discourse study as it is to grammatical investigation.

9. Consolidation

Let me mention a few obvious respects in which the previous example is oversimplified, if not simplistic. First, intonation units or attentional frames do not always coincide with clauses, and even when they do, the clause level is presumably not the only one at which discourse expectations are induced and updating occurs. Second, the situations described by successive clauses in the discourse do not necessarily form a definite or strictly corresponding chronological sequence. Obviously we have to distinguish between the sequence of presentation in discourse, on the one hand, and whatever chronological ordering there may or may not be among the presented processes within the situation being described, on the other hand. Third, a discourse is not limited to the building of a single connected structure.\(^\text{13}\) Structures can be assembled and developed in distinct channels, worlds, or mental spaces, either successively or by shifting back and forth between them. They may or may not be linked together, even when their presentation in the discourse is interspersed, as in example (19). Moreover, at any point in a discourse we can stop working on one structure and start building another.

(19) They told me I couldn’t cash a check because—Would you like some more tea? No? Just say so if you do—because I didn’t have the proper identification.
To illustrate these points, consider the discourse in (20). Here, obviously, we are not dealing with a narrative sequence, where events occur in the order recounted, but with facets of a single, stable situation. Also, between (20d) and (20e) a break occurs, where the speaker stops building one structure and starts building another. The two structures are of course connected—they are like two branches of the same tree, rather than separate trees. The break is effected in (20e) by an attentional frame, the topic phrase *Now Bill*, that is not itself a clause. Updating by this nonclausal element serves to announce that structure building is going to start in a new place.

(20)  
(a) My two children are very different.  
(b) Alice is most impressive.  
(c) She’s clearly extremely smart.  
(d) She’s also energetic.  
(e) Now Bill, he’s more ordinary.  
(f) He’s not terribly active.  
(g) He is however quite personable.

In Figure 25 I give a rough sketch of how these structures are built, step by step. We start at a point in the discourse where it has already been established that the general discussion concerns the speaker’s family. The diagram shows the successive updatings to the CDS effected by the contents of each attentional frame, as signaled by intonation. The first utterance, (20a), profiles a relationship of nonsimilarity between two children, who together constitute the clausal trajector. The next sentence is most naturally rendered with special accent on Alice: *Alice: ALICE is most impressive*. This accent singles out Alice as a local discourse topic, the reference point for a domain of knowledge in which a series of propositions are to be integrated (cf. Figure 10). This very sentence supplies the first such proposition, namely the attribution to Alice of a certain property ($p_1$), that of being impressive. The next two expressions continue building a structure in this domain by specifying additional properties ($p_2$ and $p_3$).

The following intonation group, *Now Bill*, establishes a new local topic and thus announces that structure building is going to resume in a different place, the domain of knowledge concerning Bill. That is, it induces the expectation that a proposition will be inserted there in the following frame. A series of such propositions then follow, each building on this structure by ascribing to Bill a different property.

Besides its skeletal nature, the representation in Figure 25 is clearly oversimplified in an important respect: it shows each increment of structure merely remaining in place, essentially unmodified, as further
Figure 25.

Discourse in Cognitive Grammar 179

She's also energetic.

He is however quite personable.

She's clearly extremely smart.

He's not terribly arrive.

Alice is most impressive.

He's more ordinary.

My two children are very different.

Now Bill.
increments come along. It suggests that the speaker and hearer have perfect memory of how the prior discourse has unfolded, including the specific way each successive expression construes and portrays the situation described. Most obviously, there are multiple representations of each participant, Alice and Bill, reflecting the repeated mention of them by separate nominal elements in successive utterances. I believe this kind of representation does have a certain validity. It captures the dynamicity of the unfolding discourse, showing how—through linguistic encoding—the interlocutors have accessed and constructed the complex conceptual structure under discussion. I take this as being one dimension of linguistic meaning, and representations of this sort are likely to prove essential for particular purposes (e.g., anaphora).

Undoubtedly, though, memories fade as new increments of structure come along. By the time we get to the last sentence in (20), any specific memory of the first one may be gone. While the essential content may be retained, memory of how it was presented linguistically will soon be lost. We can usefully speak of a process of consolidation, whereby the essential content is abstracted from the specifics of its linguistic presentation, e.g., by collapsing the multiple references to the same individual. For example (20), the consolidated representation would be Figure 26. It is simply an apprehension of the overall situation described in its own terms, as an integrated conceptual structure. As such, it has no record of how its components were pulled out for packaging in separate linguistic expressions, construed in particular ways, and presented in a certain sequence.

As discourse unfolds, consolidation proceeds apace. The consolidated structure continues to grow or be otherwise modified, even as the discourse structure effecting its earlier evolution fades from memory. It is the consolidated structure that we retain from earlier stages in the discourse and store in long-term memory. It too is a dimension of linguistic

![Diagram](image-url)

Figure 26.
meaning, alongside the discourse structure it abstracts away from. I would consider it a mistake to identify semantic structure with either the discourse structure or the consolidated structure individually, or with any single stage of either. The discourse structure is a kind of temporary scaffolding, put up in stages in order to construct the more permanent consolidated structure. As construction advances, and some portions are finished, the scaffolding used for those portions is taken away. Eventually the completed consolidated structure stands alone. But the linguistic meaning of any expression, along the way, consists of the section of scaffolding it represents, as well as the corresponding portions of the consolidated structure being built, both viewed in relation to the more global dynamic process in which they are embedded.

10. Conceptual structure versus constituency

A final point relates to the distinction between conceptual content and how that content is construed for linguistic purposes. The apprehension of a situation in its own terms (independently of how it might be expressed) constitutes one facet of conceptual structure and linguistic meaning. How it is coded and expressed linguistically constitutes another, which need not be isomorphic to the first. Whether in a sentence or a discourse, linguistic encoding adds meaning by virtue of the multifaceted construal it imposes.

Linguistic investigation has suffered from a tendency to confuse these two levels. A case in point is the standard analysis of complex sentences, as in example (21) for instance. Though details vary, linguists would usually assign to it some kind of syntactic tree structure with the basic constituency shown in Figure 27. This assumes the standard division of

![Figure 27](image-url)
a clause into subject and predicate, and that an object complement clause is a subconstituent within the predicate.

(21) //Alice hopes/Bill believes/Cindy left//

Note in particular that, on this account, believes Cindy left forms a syntactic constituent, as does hopes Bill believes Cindy left. These are highlighted by the first two underscores in Figure 27. Yet there is no intonational evidence to support this putative organization. If we consider a pronunciation measured enough for intonation breaks to appear, they will be as indicated by the slashes in (21), and by the lowest set of underscores in Figure 27. This apparent mismatch between grammatical and phonological structure leads to the adoption of ad hoc theoretical devices (e.g., “readjustment rules” in Chomsky and Halle 1968) to overcome the discrepancy. I would argue, however, that there is no reason to adopt the constituent structure of Figure 27 in the first place. The felt need to do so stems from the confusion of conceptual structure, organized in its own terms, and how this structure is packaged for linguistic presentation in a discourse.14

Using the basic descriptive devices of Cognitive Grammar, one can perfectly well put together a symbolic assembly for (21) which ascribes to it the constituency shown in Figure 27. This may even be appropriate in some cases. However, constituency is seen in Cognitive Grammar as being nonessential, flexible, and variable, arising as a special case out of the more fundamental phenomena of conceptual grouping, phonological grouping, and symbolization. There is no reason in this case not to posit constituency that accords with the intonational facts. As a general point, I suggest that when variation is permitted, constituency is chosen to optimize the packaging and presentation of conceptual structures in the discourse.

In Figure 28, a series of subject-verb constituents are formed at the lowest level and successively combined. There is no constituent corresponding to either believes Cindy left or hopes Bill believes Cindy left. The three intonation groups—Alice hopes, Bill believes, and Cindy left—are, however, constituents. All the proper semantic and grammatical relationships are captured, in the form of correspondences. And despite the difference in constituency, the composite semantic structure of Figure 28 is exactly the same as the one that would result from the constituency in Figure 27. It is just that the order of assembly is different, as well as the intermediate-level structures reflecting this order.

The constituency in Figure 28 directly correlates with the discourse packaging signaled by the intonational groupings. Heavily outlined boxes indicate the attentional framing, with Alice hopes, Bill believes, and Cindy
left as successive windows of attention. This is reflected in Figure 29, which shows the CDS being successively updated frame by frame. Each frame introduces and profiles a grounded process (partially schematic in the first two instances). Since they are linked together in particular ways by correspondences, a coherent conceptual structure is being built. In the second row I have shown the consolidated structures capable of emerging at each stage, basically by collapsing the representation of corresponding entities and abstracting away from construal factors such as profiling.

Comparison of Figures 28 and 29 reveals that—at the clause level and above—they contain the same conceptual structures. Moving along the top row in Figure 29, the profiled relationships at each stage correspond to the clausal constituents in Figure 28 (the successive windows of attention). Moreover, the intermediate-level composite structure Alice hopes Bill believes matches the first consolidated structure in Figure 29, and the overall composite structure matches the second consolidated structure. In other words, the grammatical constituency that best accords with the intonational evidence can be seen as correlating directly and in detail with the conceptual structures progressively assembled in the discourse, taking both consolidated and unconsolidated structures into account.

I believe this analysis captures everything that needs to be said, both semantically and grammatically. It also accords directly with the
phonological evidence, with no need for extrinsic, ad hoc devices like readjustment rules. Why, then, do linguists feel the need to posit tree structures like Figure 27, whose constituency conflicts with the overt evidence? The reason is a confusion between conceptual organization, described in its own terms, and grammatical constituency, a matter of linguistic encoding. I have already observed that these need to be clearly distinguished.

Conceptually, the component processes are nested: Cindy’s leaving is the event portrayed as what Bill believes, and this belief on his part is what Alice hopes is the case. This is just a particular configuration of mental spaces in the sense of Fauconnier (1985, 1997). It is captured in the composite structure of Figure 28, which corresponds to the final consolidated structure in Figure 29. There is no reason at all to expect this conceptual organization to be directly mirrored in grammatical constituency, which is a matter of how symbolic elements successively combine with one another to form progressively more elaborate symbolic structures. Grammar is a tool for building up to—and symbolizing—complex conceptualizations, but it is not to be identified with those conceptualizations. Constituency too is a kind of scaffolding, which enables us to reach the composite semantic structure. It is essential not to confuse the scaffolding with the structure being built.

The near identity of Figures 28 and 29 suggests the intimate commingling of discourse and grammar. This is of course a fundamental functionalist tenet, and while it has always been accepted and tacitly
accommodated in Cognitive Grammar, the discussion here may help to make the connection more tangible. In particular, the Cognitive Grammar claim that constituency is nonessential and variable can now be seen as pertaining to one side of a coin, the other side of which is discourse: how conceptual structures are packaged and presented in a discourse is itself a major factor in determining constituency; when variation is grammatically permitted, the constituency chosen on a given occasion is determined on this basis.

11. Conclusion

I have not infrequently encountered the attitude that a cognitive approach to language implies a lack of concern for its grounding in discourse and social interaction. Any such notion is both gratuitous and erroneous. For instance, a prime concern of mental space theory is how such spaces are dynamically created and accessed in ongoing discourse (Cutrer 1994; Fauconnier and Sweetser 1996; Fauconnier 1997). Moreover, an elaborate view of how linguistic structure relates to interactive language use has been central to cognitive grammar from the very outset (e.g., Langacker 1985, 1987, 1988). Here I have tried to articulate more fully how Cognitive Grammar relates to discourse. The relationship, I suggest, is natural and intrinsic. Bringing them together does not require any modifications in the framework, but is simply a matter of exploiting the potential which has been there all along. I further suggest that, for a full and explicit description of interaction and discourse, a cognitive perspective is necessary (though not sufficient). A detailed characterization of the conceptual structures being built and manipulated is as fundamental and indispensable to discourse study as it is to grammatical investigation.

Received 30 October 2000

Revision received 26 February 2001

Notes

* Author's e-mail address: rlangacker@ucsd.edu
1. See Langacker (1987, 1988, 2000). Besides its relevance to discourse, the usage-based approach (Barlow and Kemmer 2000) has important implications for both historical linguistics (Croft 2000) and acquisition studies (Tomasello 1992, 2000).
2. The representation is necessarily simplistic, with distortions stemming from the topological limitations of the two-dimensional page. (For instance, while the ground is part of the CDS, there is also a sense in which the CDS is part of the ground.)
3. These are not mutually exclusive, for we apprehend—and thereby conceptualize—our vocalizations. Moreover, the terms are to be understood quite broadly (e.g., "vocalization" subsumes gesture).

4. For fuller discussion, see Clark (1996: 263).

5. This shows once more that the specific geometry of these diagrams, constrained by the two dimensions of a page, should not be taken too seriously. There is nothing inherently problematic about the same notion, here a gesture, functioning simultaneously as a symbolizing element and part of the conception being symbolized.

6. Hence the plus frame shows a schematic proposition located in the domain of knowledge (given as an ellipse) accessible via the topic. (For discussion, see Langacker 1993, 2001.)

7. Reference here to “attentional framing” and “windows of attention” is not unrelated to Talmy’s use to such terms (1996), but pertains to a different level of organization and different channels (Figure 12). The relation between intonation units and grammar is further considered in Croft (1995).

8. For present purposes, heavily outlined boxes stand for attentional framing (they do not mark a construction’s head, or profile determinant, as they normally do in Cognitive Grammar). Heavy lines inside boxes are used for profiling (as usual).

9. This schematicity is represented by drawing the sides of the triangle with dotted lines, indicating that no specification is made concerning their relative lengths.

10. The ellipses in the minus frame for dog indicate the absence of any specification concerning the prior discourse. By contrast, the indefinite article does make a specification, albeit a negative one: it implies that the conditions for the definite article (a unique accessible instance of the type in question) fail to obtain. As an ad hoc notational device I have shown this by leaving the minus frame blank.

11. I would propose an analogous account of “pro drop” or “zero anaphora”, i.e., cases where a language permits an identified participant to remain implicit rather than being specified by a personal pronoun.

12. This commonality explains the close grammatical affinities between subjects and possessors (e.g., possessive periphrasis for the subject of a nominalization), because of which a possessor is sometimes described as the “subject” of a noun phrase. For extensive discussion of these affinities (which further extend to topics), and evidence supporting the reference point analysis of subjects, see Langacker (1999, 2001).

13. Ultimately, all the structures built are connected, if only via their mutual connection to the ground.

14. Here I am making the same point that was made at the end of Langacker (1997), but from a different perspective.

References


Chomsky, Noam and Morris Halle  
1968  

Clark, Herbert H.  
1996  

Croft, William  
1995  

2000  

Cutreri, Michelle  
1994  

Dancygier, Barbara and Eve Sweetser  
1997  

Du Bois, John W.  
1987  

Fauconnier, Gilles  
1985  

1997  

Fauconnier, Gilles and Eve Sweetser (eds.)  
1996  

Harder, Peter  
1987  

Herring, Susan C.  
1989  

Langacker, Ronald W.  
1985  

1987  

1988  

1991  

1993  

1995  

1997  

1998  
Indeterminacy in semantics and grammar. In Cifuentes Honrubia, José Luis (ed.), *Estudios de Lingüística Cognitiva II.* Alicante: Universidad de Alicante, Departamento de Filología Española, Lingüística General y Teoría de la Literatura, 649–672.

1999  

2000  

Talmy, Leonard

Tomasello, Michael


van Hoek, Karen
