

Do we need summary and sequential scanning in (Cognitive) grammar?

CRISTIANO BROCCIAS AND WILLEM B. HOLLMANN

Abstract

Cognitive Grammar postulates two modes of cognitive processing for the structuring of complex scenes, summary scanning and sequential scanning. Generally speaking, the theory is committed to basing grammatical concepts upon more general cognitive principles. In the case of summary and sequential scanning, independent evidence is lacking, but Langacker argues that the distinction should nonetheless be accepted as it buys us considerable theory-internal explanatory power. For example, dynamic prepositions, to-infinitives and participles (e.g. into, to enter, entered) are distinguished from finite and bare verbs in terms of summary vs. sequential scanning. In this paper, we try to show that various theory-internal and theory-external arguments do not seem to support the two scanning modes. In particular, we offer a detailed examination of causatives (e.g. get, make) and argue that their complementation patterns are difficult to reconcile with differences in scanning. We conclude that the status of, and need for, summary and sequential scanning in (Cognitive) grammar is doubtful, especially if grammar is approached from the viewpoint of the language learner/user. It follows also that the parcelling of some form-function mappings may prove less clear-cut than previously assumed.

Keywords: Cognitive Grammar; summary vs. sequential scanning; word classes; complementation; usage-based model

1. Cognitive abilities: summary and sequential scanning

One of the main assumptions of Langacker's Cognitive Grammar (Langacker 1987, 1991, 1999, 2002) is that general cognitive

abilities shape language.* A case in point is figure/ground segregation. Consider the examples in (1) (see Croft and Cruse 2004: 56 and Talmy 2000: 314):

- (1) a. Tom is near John.
 b. John is near Tom.
 c. The bike is near the house.
 d. ??The house is near the bike.

(1a) and (1b) seem to show that *near* is a symmetrical preposition. Its trajector and landmark, i.e. *Tom* and *John* respectively in (1a), can be swapped around as in (1b) and the resulting sentence is still perfectly acceptable. However, if the same operation is performed on (1c), the result is, under normal circumstances, an awkward sentence. The contrast in acceptability between (1c) and (1d) can be motivated on grounds of figure/ground organisation. All other things being equal, we select a smaller object (e.g. a bike rather than a house in (1c-d)) as figure, that is (roughly) as the focus of our attention (see Ungerer and Schmid 1996: Chapter 4 for an overview on figure selection; see Langacker 1987: 125, especially note 12, on the difference between figure and focus of attention).

Among the linguistically relevant cognitive abilities invoked by Cognitive Grammar are also two modes of cognitive processing for the structuring of complex events: summary scanning and sequential scanning, which pertain to processing rather than conceived or objective time (Langacker 1987: 144–145, 248–249).¹ In other words, they concern our conceptualisation of time independently of how things might occur in the “real” world. We will now move to a closer examination of the two types of scanning and their crucial importance for the “architecture” of Cognitive Grammar.

Suppose we have a complex scene, that is an event made up of different configurations (or component states) over time. For example, the verb *enter* evokes a complex scene in that it denotes (or “profiles” in Cognitive Grammar terminology) the movement of an entity (the trajector) towards a location (the landmark) resulting in the trajector’s ending up inside the landmark. The event is diagrammed schematically in Figure 1 below (the dotted lines indicate referential identity between the trajectors/landmarks of the component states; the three dots indicate that only a representative subset of all component states has been depicted).

Cognitive Grammar claims that an event (like *enter*) is scanned sequentially when the conceptualiser (i.e. the language user) views the different facets of the complex scene successively (as in a motion picture). Further, when a relation involves sequential scanning we say that the relation has a *positive temporal profile* which is indicated by the heavy TIME line in Figure 1.

Figure 1

By contrast, summary scanning obtains when the different facets of the complex scene are made available as a single Gestalt (as in a photo). Suppose a ball falls vertically. Sequential scanning for such an event is represented in Figure 2a, while the import of a summary scanning construal is depicted in Figure 2b. A short, incremental build-up phase precedes the availability of the final Gestalt.² In summary scanning the positions occupied by the ball over time are superimposed upon each other so that a holistic conceptualisation—all positions are activated simultaneously—obtains (the downward arrow in the final component is a visual aid representing downward motion explicitly). For discussion of the presence vs. absence of a temporal profile (heavy vs. light time arrow) see our description of Figure 5, below.

Figure 2

In what follows, we will present evidence against (the need to posit) the summary vs. sequential scanning distinction. Section 2 shows that the two modes of scanning are important for theory-internal reasons. The distinction between summary and sequential scanning buys the Cognitive Grammarian considerable theory-internal explanatory power in that it allows her/him to distinguish on semantic grounds between apparently synonymous verbs and prepositions (section 2.1). Further, *to*-infinitives and gerunds are also neatly classified on the basis of semantic criteria alone if the two modes of scanning are taken into account (section 2.2). However, the remaining sections of the paper will discuss issues that challenge the value of the summary vs. sequential scanning distinction. Section 3 is a case study on the issue of variation in infinitival complementation patterns with causative verbs. An adequate account can be achieved by relying on the well-established notions of semantic/syntactic binding (Givón 1980, 1990). That is, we argue that scanning would not increase the

explanatory power of our analysis and, if it were invoked, would in fact raise questions which may turn out to be very difficult or even impossible to answer. Section 4 presents further problems that the postulation of summary and sequential scanning faces. In section 5 we propose that the distinctions that scanning is intended to capture should be rethought from the viewpoint of the language learner/user rather than the professional linguist and that the parcelling of some form-function mappings may prove less clear-cut than previously assumed. In section 6, finally, we take sides with Taylor (2003a, b), as regards the potentially conflicting dual nature of Cognitive Grammar as an epistemic model and a usage-based model. We conclude the paper by showing that Langacker's (2006, p.c.) position on scanning may actually be considerably more sophisticated than his published discussions would seem to suggest, but we argue that usage should be brought to the fore even more than is presently the case.

2. Why scanning is needed in Cognitive Grammar

In this section, we will argue that the two types of scanning are needed first and foremost for theory-internal reasons, a point which is recognised explicitly by Langacker himself but is usually ignored in the literature (e.g. Croft and Cruse 2004: 53–54, Taylor 2002). The problem here is that theory-internal coherence does not necessarily guarantee a psychologically plausible linguistic theory. In order to validate the distinction between the two scanning modes external evidence would be desirable — as is indeed commendably provided for other basic theoretical notions in Cognitive Grammar, such as figure/ground segregation (cf. section 1, above).

2.1. *Enter vs. into*

Among the merits of the postulation of sequential and summary scanning is the fact that it allows for the distinction between *processes* and *atemporal relations*. In order to appreciate this point fully, readers who are only partially familiar with the theory should consider that Cognitive Grammar defines word classes semantically rather than distributionally. Langacker thus assumes that any element (including morphemes such as *-ing*, infinitive marker *to*, etc., which in more traditional theories such as Generative Grammar are considered to have grammatical rather than semantic import) is meaningful. Every linguistic expression is analysed as a pairing of meaning (its semantic pole) and form (its phonological pole).

The Cognitive Grammar analysis of (the semantic pole of) word classes is summarised diagrammatically in Figure 3 (shaded boxes indicate concepts that are possible semantic poles for word classes; for example, entities, unlike things, do not correspond to any particular word class). We start with the very general category of *entities*, which can be divided into *things* and *relations*. A thing, a technical term in Cognitive Grammar, is a set of interconnected entities and is the semantic pole of the noun class. For example, the noun *team* (see Langacker 1987: 197) profiles a set of entities (represented in Figure 4a as the dashed boxes connected to each other) rather than singling out any constitutive member. The emboldened circle visually represents such a set in Figure 4a. Relations, by contrast, profile connections between entities (as well as the entities themselves), as is shown for two arbitrary entities e_1 and e_2 in Figure 4b. Relations can be either *processes* or *atemporal relations*. The former involve sequential scanning, that is, they are relations with a positive temporal profile or, to put it differently, they are relations scanned as a sequence of configurations. Processes, depicted schematically in Figure 4c, constitute the semantic pole of verbs (e.g. *enter*). Atemporal relations have a “null” temporal profile and come in two types: *stative relations* and *complex atemporal relations*. Stative relations involve a single, stable configuration through time and correspond to the semantic pole of adjectives (i.e. stative relations whose trajector is a noun and whose landmark is a region on a scale), adverbs (i.e. stative relations whose trajector is a process and whose landmark is a region along a scale), and stative prepositions (such as *in*, as opposed to *into*). Complex atemporal relations (e.g. *into*) are made up of more than one configuration over time but such configurations or facets are scanned in summary fashion.

Figure 3

Figure 4

Langacker (1992: 290) claims that the contrast between sequential and summary scanning is needed, among other things, to distinguish between the verb *enter* and the dynamic preposition *into*.³ Remember that since word classes are defined semantically (rather than distributionally) and since the scenarios evoked by *enter* and the motion preposition *into* are intuitively similar (i.e. one entity moves towards another, which is fixed, and ends up

inside it), we need a *semantic* criterion to distinguish between the two.⁴ Cognitive Grammar suggests that, whereas *enter* profiles a sequentially scanned relation, *into* activates summary scanning, i.e. the various facets of the denoted event are made available as a single Gestalt. A representation of *into* along the lines of Figure 1 and Figure 2b is offered in Figure 5.

Figure 5

There are two obvious (related) differences between Figure 1 (i.e. the schema for *enter*) and Figure 5. First, each configuration in the latter diagram keeps track of the preceding ones as the superimposed circles are intended to show. Second, the TIME arrow is not heavy because sequential scanning is suspended: all facets of the motion event are said to be made simultaneously available, i.e. after the build-up phase.

2.2. Infinitives

The two types of scanning are also recruited to provide a semantic characterisation of bare infinitives (e.g. *enter*) vs. *to*-infinitives (*to enter*) vs. gerunds or *-ing* forms (e.g. *entering*). The rationale is always that differences in form must imply differences in meaning.

We have already pointed out that bare infinitives are relations scanned sequentially. As for *to*-infinitives, Langacker observes that

they are not in fact verbs by my definition — instead they designate atemporal relations. They nevertheless derive from verbs [...] More specifically, the process designated by the verb stem functions as the base for the infinitival [...] predication overall. The semantic value of the derivational morphology (*to* (...)) resides in the effect it has on the process introduced by the stem: [it] imposes its atemporal profile on the processual base provided by the stem. (Langacker 2002: 82)⁵

Finally, *-ing* forms are also said to be complex atemporal relations. Still, “they differ from infinitivals by construing the component states of the base process as effectively homogenous and profiling only a representative series of the states” Langacker (1987: 249, footnote 3).

Although Langacker’s theory nicely captures the similarities existing between *to*-infinitives (and *-ing* forms) and adjectives, other researchers have offered alternative analyses and raised problems for Langacker’s model. Duffley (2003) remarks that the

contrast in (2) is left unexplained in Langacker's analysis. While (2a) means that John didn't forget to lock the door but might have no recollection of doing so, (2b) means that he was aware of having locked the door. The difference between the two sentences obtains naturally if *to* infinitives are associated with some purposive meaning due to the possible (ultimately etymological) relation between infinitival *to* and (motion) preposition *to*.⁶

- (2) a. John remembered to lock the door.
b. John remembered locking the door.

A similar point is made by Pelyvás (2006). He claims that *-ing* forms are not necessarily scanned summarily and, in particular, that "since *to* highlights path and in an abstract sense intention and potentiality, its conceptual structure may not be incompatible with a similar [i.e. non-summary scanning, CB/WBH] analysis". He also remarks (note 11) that his suggestion may be supported by the fact that modal *ought* and *should* are very similar in meaning although the former takes a *to*-infinitive and the latter a bare infinitive (one might, however, argue that *to* is becoming fused with *ought* and, hence, no real syntactic difference obtains between the two).

Duffley (2006) offers yet more criticism, in observing for instance that in the sentences reproduced below the *-ing* form seems more likely to feature sequential than summary scanning (i.e. we intuitively play the events of strolling and tearing in our minds as "motion pictures" rather than perceiving them as "photos", see section 1 above).⁷

- (3) The woman strolling down the beach is my mother.
(4) I found my little brother tearing my photo album to pieces in my bedroom.

Recent research into infinitival complementation by Hamawand (2002, 2003a, 2003b) takes the existence of and distinction between the two scanning modes for granted.⁸ By contrast, in his analysis of *-ing* complements, Egan (in preparation) suggests the possibility that there may be another processing mode, which he dubs "recursive scanning". He argues that some *-ing* patterns, like (5) below, cannot be described adequately by invoking either summary or sequential scanning.

- (5) I recall her pacing the sitting-room while I am doing my homework, pausing every so often to stand at one of the windows and look down into the busy street below. (BNC HD7 1331)⁹

In Langacker's theory, an *-ing* process implies that its component states are always construed as homogenous, i.e. indistinguishable from each other. Egan observes that the *-ing* event described in (5) cannot be considered homogenous or at least invites a non-homogeneous construal because, for example, pauses in the *pacing* event are explicitly mentioned. He also seems to suggest that sequential scanning cannot be resorted to as the scanning mode for the *-ing* process because only the central portion of the *pacing* event, rather than its starting and end points as well, is profiled. (Incidentally, it should be pointed out that Egan, by conceding the possibility that *-ing* processes may be compatible with sequential scanning, is actually relaxing Langacker's characterisation of *-ing* processes. In Langacker's theory, *-ing* processes always involve summary scanning alone.) Egan proposes that the type of scanning involved in the *-ing* process in (5) is recursive scanning: we are supposed to go through the event in question repeatedly without kick-starting it.

Since Egan's research is still in progress, it is premature to discuss it any further, but it should be clear that not all Cognitive Grammarians recognise that the two scanning modes may be adequate or sufficient to capture the interpretation of complementation patterns.

Finally, it should also be observed that there seems to be disagreement in the case of nominals. Remember (see section 2.1) that nominals always involve summary scanning (since they do not have a positive temporal profile by definition). Matlock (2004b), however, invokes summary and sequential scanning to explain the differences in acceptability of nominals in fictive motion sentences, i.e. sentences like *A trail goes through the desert*, where a motion verb is used to describe a static scene (see e.g. Talmy 2000). She suggests that the difference in acceptability between (6a) and (6b) may depend on whether the subject NP's referent is scanned sequentially (see (6b)) or not (see (6a)).

- (6) a. ??The cell phone goes from the cup to the book.
(phone on desk)
b. The cell phone goes from the cup to the book.

(phone in ad on billboard)

Matlock argues that a cell phone is not usually scanned sequentially (even if it is unusually long) because “a coherent whole can be obtained with just one glance” (2004b: 228), thereby explaining the oddness of (6a). Since (a sufficiently large amount of) spatial extension is crucial to the acceptability of fictive motion sentences, only reference to the cell phone on a large billboard can result in an acceptable sentence. Although the confirmation of this hypothesis is left to future research, we observe again that the notions of summary and sequential scanning are not always used in accordance with Langacker’s theory.

An additional source of potential confusion is the fact that Langacker himself distinguishes various types of scanning other than the summary and sequential modes. First of all, scanning is sometimes used in the very general sense of the operation connecting the standard and the target in acts of comparison (see e.g. Langacker 1987: 102). Second, in discussing spatial bounding and shape, Langacker uses the terms “field scanning”, “expanse scanning”, and “periphery scanning” (the interested reader is referred to 1987: Ch. 5, section 4). Third, Langacker’s analysis of what is widely known as fictive motion (see e.g. Matlock 2004a, b; Talmy 1983) involves a contrast between perfective and imperfective virtual motion, see, respectively, examples (7–8) below (from Langacker 2005: 175):

- (7) The path is rising quickly as we climb.
- (8) The path rises quickly near the top.

For these two sentences Langacker provides the following schematic characterisations:

Figure 6

We note that Figure 6a is similar to the visual representation of sequential scanning, and that Figure 6b is similar to summary scanning (see Figure 2, above). However, Langacker does not make any overt connection with these two scanning modes. Moreover, he does not state explicitly what the diagrams portray: the verb *rising/rises* alone or the clause as a whole. This is problematic because in terms of summary and sequential scanning we would expect *rising* to invoke summary scanning, and *rises* to

be scanned sequentially. In our view, which is agnostic as to the status of summary and sequential scanning, Langacker nonetheless makes an important point in analysing example (7) as implying a local view, and example (8) as “tak[ing] a global view” of the situation (2005: 176). This is in line with the standard view that the progressive imposes an internal view on an event in that the temporal boundaries of the event are ignored (see e.g. Comrie 1976: 4; for an earlier expression of the same view see e.g. Charleston 1960: 162).¹⁰

Finally, it could be claimed that summary scanning is also needed to distinguish nominalisations of verbs, see e.g. *Something exploded vs. There was an explosion* (Langacker 2002: 98).¹¹ The exploding event in the latter sentence is said to be scanned summarily. We agree with Langacker that “[n]ominalizing a verb necessarily endows it with the conceptual properties characteristic of nouns” (2002: 98) but we would dissociate reification of events from the question of scanning. This is also done, for example, by Croft (2001: 88), who analyses action nominals as action words used in the discourse prepositional act of reference, i.e. without invoking summary scanning as defined by Langacker.¹²

2.3. *Theory-internal reasons*

Langacker himself is aware of the somewhat speculative nature of his analysis (1987: 235–254), see also his more recent reiteration (1999: 223) that the distinction between summary and sequential scanning has not been verified experimentally yet. Nonetheless, Langacker (1987) defends his analysis by claiming that summary and sequential scanning are primarily needed in order to achieve theory-internal coherence. The relevant passage is worth quoting (almost) *in toto*:

A hard-nosed linguist will doubtless ask for evidence to support these claims. How can one prove that the conception of a process (hence the meaning of every verb) requires sequential scanning [...]? The request for justification is certainly legitimate, but we must take some care that the form of the request does not embody methodologically unreasonable expectations. In particular, one cannot reasonably expect or demand the existence of direct empirical evidence that bears on this question alone considered *in isolation from the overall descriptive context in which the analysis of processes is embedded* [emphasis ours]: I can no more substantiate the claim that verbs imply sequential scanning—*directly, and without regard to how the total descriptive system meshes together*

[emphasis ours]—than the proponent of a more fashionable model can prove that movement rules leave traces without explicating the function of these constructs as part of a much larger theoretical and descriptive framework. The absence of direct and conclusive empirical support is unfortunate, but no linguistic theory can provide such motivation for all its constructs taken individually. (Langacker 1987: 253)

Indeed, without recourse to summary and sequential scanning, if we only relied on semantic considerations independently of distributional facts, it would be less straightforward to distinguish between, for example, the motion preposition *into* and the bare infinitive *enter* (see section 2.1 above, but also section 5 below for a semantic account that does take distribution into account).

We would like to argue that this is not a sufficient reason to accept the *a priori* existence of the two types of scanning without some (direct or indirect) evidence. In order to achieve external coherence, i.e. a psychologically plausible linguistic theory, all linguistically relevant cognitive abilities postulated by Cognitive Grammar must be supported by (direct or indirect) independent evidence, or at the very least be in principle amenable to experimental verification. In what follows, we try to show that such support is lacking at present and that various pieces of evidence that *are* available to us, both theory-external and internal, render the postulated distinction less than obviously necessary, and perhaps even problematic.

3. A case study: Variation in causative verbs

The discussion so far has not referred to much linguistic data, but the linguistic facts themselves suggest that the sequential vs. summary scanning distinction is problematic — at least the way in which Langacker has related this distinction to the data. Specifically, the suggestion that bare infinitives (usually called *stems* in Langacker's work) feature sequential scanning while marked (i.e. *to-*) infinitives involve summary scanning is difficult to maintain in the face of verb (or construction) classes where there is variation in complementation patterns (see also Pelyvás's 2006 observations mentioned in section 2.2). Let us consider the variation in causative verbs here:¹³

- (9) He had his secretary order some coffee, then closed the door and sat down behind his desk. (BNC ECK 2589)

- (10) The police got him to confess to the crime. (BNC HXG 799)

On Langacker's account, in causative constructions such as *have* or *make* the bare infinitive caused event is processed by means of sequential scanning, while in *get*, *cause*, *force* or *persuade* the lower clause event is accessed through summary scanning. It is difficult to accept Langacker's proposal especially in the light of accounts of the meaning and form of causatives that are less resistant to testing against empirical facts. Givón's (1980, 1990) discussion, based on the crosslinguistically valid notion of binding, is a case in point. Binding relates not only to causatives but also to "modality verbs" (*want*, *succeed*, *start*, etc.) and "cognition-utterance verbs" (*think*, *say*, etc.) (Givón 1980: 333). It is defined as the extent to which the matrix and lower clause events are coded and conceptualised as a single, integrated event. Thus, binding has a syntactic and a semantic dimension. They correlate as follows: "The higher a verb is on the [semantic] binding scale, the less would its complement tend to be syntactically coded as an independent/main clause" (Givón 1980: 337). This correlation is iconically motivated by the proximity principle: linguistic distance may be used to mirror conceptual distance (e.g. Haiman 1985: 102–147). Regarding the use of subordinating complementisers such as *that*, or indeed *to*, Givón writes:

All other things being equal, the use of a subordinating morpheme which neatly separates the main clause from its complement clause is a coding acknowledgement that the two clauses are semantically still independent of each other, at least to some extent. (Givón 1980: 371)

Let us see how syntactic and semantic binding are defined. Syntactic binding is the most straightforward. The cross-linguistic facts suggest that there are three aspects to coding as an independent clause:

- (i) The degree to which the agent/subject/topic marking of the embedded-clause agent/subject reflects the marking in independent main clauses.
- (ii) The degree to which independent-clause tense-aspect-modality marking of the verb is preserved in the embedded clause.
- (iii) The presence or degree-of-presence of predicate-raising of

the complement verb into the main verb; i.e. the degree to which the complement verb is lexicalized as one word with the main verb (Givón 1980: 337).

Infinitival complements compared to e.g. *that*-clauses represent rather extreme cases of the reduction of tense-aspect-modality marking (Givón 1980: 337)—bare infinitives even more so than *to*-infinitives. The degree of semantic integration with bare infinitives should therefore be at least as high as with *to*-infinitives.

Semantic binding is less straightforward because it depends on the class of complement taking verbs, i.e. causatives vs. modality verbs vs. cognition-utterance verbs. For causatives, Givón originally proposed two properties: (i) intended vs. unintended causation and (ii) direct vs. mediated causation, i.e. absence or presence of an intermediary party in the causal chain (1980: 336). Givón (1990: 520–526) adds two further parameters: unity of space and unity of time. (Due to the fact that these two factors hang together, also with direct vs. mediated causation, these three properties represent the typological notion of directness.) The first value in each pair represents a higher degree of semantic binding. (For modality and cognition-utterance verbs the relevant factors involve the extent to which the higher clause subject makes a stronger or weaker attempt to achieve the lower clause event, and the degree to which they are committed to its success or truth (see Givón 1980: 342–347 and *passim*)).

In connection with the correlation between syntactic and semantic binding it is important to note that the mapping between the two is not a simple one, in the sense that the semantic binding scale does not allow any absolute predictions concerning syntactic integration—only relative predictions:

If a point on the semantic hierarchy of binding is coded by a certain syntactic coding device, then a semantically higher point cannot be coded by a syntactically lower point. Rather, it will be coded either by the same coding point, or by a higher coding point on the syntactic coding scale. (Givón 1980: 370)

Similar suggestions concerning the form-function mapping in causatives and other complex predicates have been made by other authors, such as Cristofaro (2003), Dixon (2000), Duffley (1992), Fillmore (1972), Fischer (1992, 1995, 1997a, 1997b, 1997c, 2000), Jackendoff (1972), Mittwoch (1990), Wierzbicka (1975).

Hollmann (2003: Ch. 5, 2005) surveys the literature and argues that Givón's factors should be supplemented by three more:

- (iv) presence vs. absence of a sphere of control frame (i.e. of the causer over the causee)
- (v) causation type, according to Talmy's (1976, 1985, 1988, 2000) four-way typology: affective, physical, volitional and inducive causation (see also Croft 1991: 167)
- (vi) punctuality of the causing event

The aggregate furnished by these six parameters is referred to as the extended binding hierarchy. As for (iv), the suggestion is that a causative situation where the causer inherently controls the causee (e.g. socially or physically) is easier to conceptualise as a single integrated event than a situation where this is not the case.¹⁴ *Have* describes causation against the background of a control frame:

- (11) John *had* his daughter *tidy* her bedroom.
- (12) ?Five-year-old Alice *had* her father *tidy* her bedroom.

Force—which significantly takes a *to*-infinitive—is the opposite: the reason why the causer resorts to the use of force is that there is no implicit relation of superiority between them and the causee.

To see how Talmy's causation types are related to binding it is necessary to explain this classification a little. The basic hypothesis is that causers and causees are analysed according to their animacy, i.e. animate/human vs. inanimate—or mental vs. physical, as in the following diagram from Croft (1991: 167):

Figure 7

The arrows in this diagram represent the manipulative interaction between causer and causee. Verhagen and Kemmer argue that there are differences in the degree of semantic binding (which they refer to as *directness*, notwithstanding the more common use in typology referred to above) between the four types:

An obviously important aspect of this model of causation types is the very marked *asymmetry* between entities with a mental dimension (animates) vs. those that are merely physical. Animates can only act on animates via the intervening physical world, i.e. the model implies that one cannot reach into another person's mind and *directly* cause him or her to do, feel,

or think something. Physical entities are taken to act directly on other things; hence the straight arrows in the diagram in Fig. [7], vs. the very bent arrow for mental-on-mental causation, and the slightly bent one for mental-on-physical. (Verhagen and Kemmer 1997: 71)

In other words, Verhagen and Kemmer suggest a (partial) ordering of causation types such that physical and affective causation are easiest to conceptualise as a single integrated event, while volitional causation is seen as less integrated, and inductive causation features the loosest bond between causing and caused events:

physical, affective<volitional<inductive

As for punctuality, finally, the hypothesis is that this has an impact on binding in that an instantaneous causing event is construed as more integrated with the caused event than is a causing event that is seen as being stretched out over an extended period of time.

Using data from the FLOB corpus¹⁵ Hollmann (2003, 2005) shows that this aggregate of factors adequately “predicts” the complementation pattern (bare vs. marked infinitive) in *cause*, *force*, *have*, *get*, *make* and *persuade*. All tokens of these causatives were collected, and scored for the various semantic binding parameters. While there was not always complete consistency across instances in terms of their semantics, clear patterns nonetheless did emerge. *Get* and *persuade*, for example, were found to describe inductive causation most of the time but not always. Causation is of the inductive type in, respectively, 93 and 95 per cent of the cases, which warrants the conclusion that these constructions are prototypically associated with this type. The overall results of the analysis are presented in Table 1 below. Directness is used in the typologist’s sense of the term, i.e. it subsumes unity of time, unity of space, and absence vs. presence of an intermediary party. The parameter relationality is a combination of intendedness, sphere of control and causation type. The reason for grouping these together also lies in their clear interrelatedness. Consider for example that in order for causation to be intended, the causer must be human, or at least animate. They must also be human if the causee falls within their sphere of control, as inanimate objects cannot meaningfully be said to control anything.

Table 1

The analysis presented in this table clearly supports the idea that semantic binding motivates syntactic binding: *have* and *make*, the only causatives here that take a bare infinitive, are also the only predicates with the maximal scores across all three semantic binding parameters.

As a final note, we observe that a scanning-based account of complementation in causatives raises even more questions in relation to the historical development of these constructions. In Middle English there was a considerable amount of variation in infinitival complementation in causatives, see e.g. the examples of *make*, below, both of which were obtained from the earliest Middle English subperiod (1150–1250) of the Helsinki Corpus.¹⁶ More instances of this variation can be found in e.g. Visser (1973: 2256–2284), the *Oxford English Dictionary (OED)* and Hollmann (2003).

- (13) Sunnedei aras ure drihten from deðe to liue. and makede arisen mid him alle þa þet him efden er ihersumed. (HM1 IR HOM LAMB14 141)
 ‘On Sunday Our Lord arose from death to life. And he made arise with him all those who had formerly obeyed him.’
- (14) lo þe sweoke hu he walde makien hire aleast to leapen in to prude. (HM1 IR RELT ANCR 121)
 ‘Lo the traitor, how he wanted to make her at last jump into pride.’

After Middle English a regulation process set in, whereby complementation in causatives became more or less fixed. (This process was completed in c.1800.)

Langacker’s characterisation of bare and marked infinitives in terms of the two scanning modes is made with reference to examples from Present-day English. One might wish to argue that in older varieties this distinction did not obtain, but that immediately raises the question as to how the present-day situation came about. If we assume, by contrast, that the two infinitival modes did correspond to different scanning modes in 18th century English and before, then the implication is that speakers had a choice (in some sense) in their construal of the lower clause event, i.e. as scanned summarily or sequentially. The problem, once

again, is that this claim cannot be tested. Moreover, given that by around 1800 the variation in complementation had pretty much disappeared, one would have to explain why this freedom of construal was lost, taking account of the fact that some causatives came to be associated with lower clause sequential scanning (*have*, *make*), while some other causatives (*cause*, *force*, *get*, *persuade*) became linked to summary scanning.

Hollmann (2003: Chapter 5) offers a possible explanation that makes no reference to a difference between scanning modes. Whilst this is not the place to go into a lot of diachronic detail we note that one factor in the regulation process may have been the relative frequency of the constructions. It is widely accepted in the usage-based model and grammaticalisation theory (see e.g. Bybee and Scheibman 1999), that high token frequency constructions will tend to get reduced more than low frequency ones. Soon after its rise in the Middle English period causative *make* with an infinitive became the most frequent causative, which may help explain why it ended up with the relatively compact bare infinitival pattern as against the longer *to*-infinitive complement. (For periphrastic causative *have*, which has always been less frequent, the explanation must rely more on the semantics of the construction.)

4. Further challenges

The previous section has shown that complementation patterns with causative verbs can be explained satisfactorily without resorting to the notion of scanning. In fact, if summary and sequential scanning were appealed to, the analysis would raise various questions potentially very difficult to answer. In order to show that the postulation of summary and sequential scanning has important repercussions not only for specific cases like complementation patterns but also for the whole theory of Cognitive Grammar, we now turn to further challenges, both theory-internal and theory-external.

Let us first reiterate the point made in section 2, that, to the best of our knowledge, no psycholinguistic evidence is available which confirms the existence of the two types of scanning (nor is there evidence for Egan's (in preparation) third mode, of course, since it has just been proposed). For example, experiments carried out by Catherine Harris of Boston University proved inconclusive (Harris 2002, p.c.).¹⁷ Matlock's research on fictive motion, mentioned in section 2.2, may be a potential source for evidence bearing on the distinction between the two scanning modes. Here, we would like

to make the more general point that, although Matlock (see e.g. 2004a, b; Matlock et al. 2005) demonstrates that mentally simulated motion is involved in fictive motion processing, she observes that at the present stage of our knowledge the question of how we actually simulate motion while processing fictive motion sentences cannot be answered. In fact, she suggests three possibilities: (a) we activate a static linear or path-like model and later simulate movement along that path; (b) the path representation is built gradually; (c) we do not simulate motion, as in (a), or scanning, as in (b), but a state change, i.e. we imagine a series of points along the path, each a few milliseconds after the other. To our mind, (a) might correspond to Langacker's summary scanning and (b) to sequential scanning, while (c) does not seem to be contemplated in Langacker's theory. Importantly, not only is it impossible at present to distinguish among the three options experimentally but Matlock also suggests that people may actually combine the three types.

To throw the lack of experimental evidence into relief, consider that other linguistically relevant cognitive processes, such as prototype and schema-based categorisation and figure-ground segregation, have been well-documented *prior to* their use in Cognitive Grammar theorising. The summary vs. sequential scanning distinction, by contrast, almost twenty years after it was postulated, still awaits experimental confirmation. This is especially striking given the status of these modes as *basic* cognitive operations (i.e. they are not regarded as being reducible to a complex of more basic cognitive operations, at least in our understanding of Langacker's theory). It is perhaps not surprising that summary and sequential scanning have attracted (sometimes strong) criticism before, for instance by Francis (2000), who characterises these modes as "highly esoteric concepts for which there could be no counterexamples" (*ibid*: 100, also cited in Taylor 2002: 516).

We would like to point out that the postulation of summary and sequential scanning may be an instantiation of the *post hoc propter hoc* fallacy. Sequential scanning in grammar implies that an element X can be inflected (e.g. *enter* can be inflected for person/tense as in (*she*) *enters*). But, if we do not provide some language-independent evidence for its existence, we know that X evokes sequential scanning only from the fact that X can be inflected. Further, if tense inflections are a valid criterion for establishing whether a form involves sequential scanning, what

does that imply for languages where verbs are not inflected for tense, e.g. Thai, Vietnamese and Malay? One suspects that the summary vs. sequential scanning distinction may be hard to maintain in the face of evidence from these languages.

The issue of psychological plausibility should also be addressed. Consider, for example, Langacker's (1991: 199) analysis of the verbal group *have been being followed* of sentence (15), which is reproduced in (16):

(15) I may very well have been being followed.

(16) (*have* (PERF₄ (*be*₁ (-*ing* (*be*₂ (PERF₃ (V)))))))

In other words, we start with the temporal relation symbolised by *follow* (i.e. V), which by definition involves sequential scanning, and combine it with PERF₃, thus obtaining an atemporal relation, i.e. the participial form *followed*, which requires summary scanning (for simplicity's sake we will not offer a detailed description of the semantic pole of the various (indexed) morphemes contained in (16) since this is not pivotal to our argument). *Followed* combines with *be*₂ to derive the temporal relation *be followed*. *Be followed* merges with the morpheme *-ing*, which atemporalises the relevant relation again. Sequential scanning obtains when *being followed* fuses with *be*₁, thus giving rise to *be being followed*. Next, PERF₄ imposes summary scanning on the resulting predicate *been being followed*. Finally, the relation is re-temporalised thanks to perfective *have*. In sum, verb groups like those in (15) originate from the cyclical application of summary and sequential scanning, resulting in the *compositional path* summarised in (16). One may wonder what the psychological reality of such cycles is: do speakers really alternate between the two types of scanning? If much in grammar is accessed as a unit (i.e. automatically or without much constructive effort), as Cognitive Grammar itself claims, speakers could/should have access to the schematic structure(s) underlining instantiations like the one in (15) without having to go through the "generative" procedure elucidated in (16), i.e. without constantly having to "switch" summary and sequential scanning construal on and off. In fact, Langacker himself suggests that the pattern in (16) (or some of its subpatterns) may be stored as a unit (1991: 227). If we interpret this as an indication that (some?) speakers do not need to run through the entire cycle of alternating scanning modes, instead

selecting the appropriate scanning mode immediately, the question still remains as to how they acquire this knowledge in the first place if not through instances of this kind of oscillation (see also the point above on the *post hoc propter hoc* fallacy).¹⁸

It is worth dwelling further on the analysis of (15). As it stands in (16), it is also not clear, when the scanning mode oscillation applies, whether (16) is intended to capture processing on the part of either the speaker or the hearer (or both). Since Langacker's theory is a *cognitive* theory of language, any explanation Cognitive Grammar offers must be compatible with psycholinguistic evidence. Although processing is a highly contentious issue, particularly in the case of speech production, we will try to argue that in either case the evidence available to us casts doubts on the appropriateness of Langacker's analysis for (15) as a *cognitively* plausible *processing* representation.

Let us suppose that the analysis offered in (16) applies to processing on the part of the hearer, i.e. comprehension. This would mean that we, as hearers, would begin processing the verbal group at the penultimate morpheme (i.e. *follow*). We would then move to PERF₃. But if we have not processed any of the preceding forms at that stage, it is not clear how we could analyse *followed* as a participle — scanned summarily —, as it is syncretic with the simple past — which involves sequential scanning. A more plausible analysis of the way hearers process sentences like this starts from the widespread recognition, in cognitive psychology, of the so-called *immediacy of interpretation*, i.e. the notion that sentence processing is done to a large extent by assigning syntactic/semantic interpretations to words as they come in, as opposed to only at the end of the sentence. Immediacy of interpretation is supported by important experiments carried out in the 1970s and 1980s, see e.g. the study of eye movements by Just and Carpenter (1980), or the computational simulation experiments by Reddy (1975, 1980). For an overview of the literature the reader is referred to e.g. Barsalou (1992: 243–244 and *passim*) or Anderson (1995: 383–385). With reference to example (15), the psycholinguistic evidence suggests that already after hearing the very first auxiliary, *may*, the hearer will hypothesise that this is the beginning of a verb phrase (see also Barsalou 1992: 234 on identifying constituent types in real-time language processing). The subsequent auxiliaries and finally the lexical verb *followed* will confirm this hypothesis, and in identifying the latter verb as a past participle rather than a simple past, the hearer is obviously helped

considerably by the knowledge that the preceding verb, *being*, may be a passive auxiliary, which requires a past participle to complement it. Thus, if instead of Langacker's perspective we take the view that is congruent with the experimental evidence available on speech comprehension, the syncretism problem does not arise.

Having shown that Langacker's analysis of (15) does not seem to be consonant with experimental evidence on comprehension, we may hypothesise that it must apply to production. Due to methodological problems we have far less experimental evidence bearing on language production than we do for comprehension. Much of the evidence we do have involves speech errors (e.g. Fromkin 1971, 1973; Garrett 1975, 1980, 1988). Despite its indirect nature, this evidence clearly points in certain directions that suggest that Langacker's compositional path for sentences such as (15), above, cannot be assumed for the production dimension of usage either. It is not appropriate here to give a full overview of the literature on language production (but see e.g. Clark and Clark 1977: Chapters 6–7; Dell 1986; Levelt 1989).¹⁹ Simplifying matters considerably, we can say that, disregarding some differences between the various models, most psycholinguists agree that language production involves several stages or levels, moving from the purely conceptual (propositional) level, via semantics and syntax, to phonology and ultimately phonetics.²⁰ And, importantly, neither on the conceptual level nor on the semantic and syntactic levels does it seem to be the case that speakers start out at the lowest level of constituency, and then work their way up, step by step, in the tree or hierarchy. With reference to example (15), when on the conceptual level we form the proposition we intend to convey there is no reason why we should start with the event corresponding to *follow*. Instead, the proposition is likely to be a single Gestalt involving (some modalised version of) someone following the speaker. As regards the stages of the process where the message is converted into language, the consensus opinion among psycholinguists is that high-level schemas—such as in this case the transitive frame of *follow* and of the English passive construction—are evoked very early, and that the semantic representations corresponding to the entities and relations of the propositions are slotted into place. These semantic representations are associated with words. Generally, lexical words are assumed to be inserted into the structures/schemas earlier than function words. This might seem to provide support for Langacker's suggestion concerning example

(15), that the compositional path starts at the lexical verb *follow*. However, this support is only partial at best, as it is not clear why the speaker should subsequently insert the auxiliaries in the exact order Langacker stipulates. Moreover, the fact that a high-level schema such as that of the English passive is activated relatively early means that the hierarchy is not assembled in the strictly bottom-up manner of the hypothesised compositional path given in (16). We thus conclude that from the point of view of the speaker, too, compositional paths involve top-down processing as well.

Of course, one could claim that the compositional path in (16) is a convenient representation of the interaction between the two scanning modes and does not reflect how, for example, speakers actually build up (or decode) sentences like (15) – see above on Langacker’s (1991: 227) suggestion that (parts of) (16) may have unit status and see e.g. Langacker (1999: Ch. 5) for hints as to non-strictly bottom-up compositionality which may be consonant with the psycholinguistic evidence mentioned here (but note that such observations seem to apply only to production rather than comprehension). However, both unit status and non-strictly bottom-up compositionality, presumably two sides of the same coin, are obviously compatible with a model which dispenses with the existence of the two scanning modes. In fact, unit status and non-strictly bottom-up compositionality “hide” scanning oscillations, making it all the more difficult to prove their existence and their scope of application (*vis-à-vis* unit status). As was the case with causative verbs (see section 3), the postulation of the two modes does not render our explanations more effective but, rather, adds a conceptual dimension for which evidence seems to be very elusive (at least at present).

Indeed, the issue of the postulation of summary and sequential scanning may blur the distinction between language as an object of investigation on the part of the professional linguist and language as a cognitive representation in the speaker’s mind (cf. Croft 1988; Sandra and Rice 1995), although Cognitive Grammar aims at developing a psychologically plausible linguistic theory. This point is explicitly acknowledged by Taylor:

As was the case with vowels and consonants, there is an important sense in which the categories of adjective and noun (and indeed the other word classes) must be understood with respect to the constructional schemas in which they occur (Croft 1999). This is not to deny the possibility of entertaining construction-independent characterizations of the word-

classes, in terms of the nature of the concepts that the words designate, for example (Chapter 9). Ultimately, however, a word class emerges as a function of its role within a constructional schema. (Taylor 2002: 563)

The quotation, which is consonant with recent, usage-based research into the acquisition of language by children (see e.g. Tomasello 2003, who shows that language acquisition relies on constructions), brings to the fore the very notion of distributional evidence which Langacker had not used to define word classes (see section 2).

More generally, the combination of the desire to see all linguistic elements as meaningful and the recognition of entrenchment (i.e. the view that because of repetition much in language is accessed automatically), constitutes a potentially problematic duality in Cognitive Grammar. On the one hand, Cognitive Grammar is a semiotic model (see also Taylor 2003b) where all elements are said to be meaningful. This requires maximum parcelling of meaning. On the other hand, Langacker presents Cognitive Grammar as a usage-based model (or corpus model, see Taylor 2003a, b). On this view, grammar is emergent: it emerges out of concrete forms which an individual is exposed to and can manipulate. The issue here is that entrenchment may sometimes be in inverse proportion to analysability (see e.g. Croft and Cruse 2004: Ch.9 on the relation between conventionality and compositionality). The semiotic perspective is linguist-oriented (and possibly diachronically oriented), in the sense that the linguist motivates the (coming into) existence of a given structure. The usage-based view is language user-oriented and may not require detailed representations like the one in (16) above.

5. On *enter* and *into* again

If we do not accept the existence of the two types of scanning or, more cautiously, remain agnostic about them and recognise the centrality of distributional facts, we can still solve the nagging problem of distinguishing between the verb *enter* and the dynamic preposition *into*. Distinctions such as those between processes (e.g. *enter*) and atemporal relations (e.g. *into*) can be regarded first and foremost as by-products of distributional facts, i.e. of grammar as a usage-based model. The fact that *enter* is a verb and *into* is a preposition does not (necessarily) either stem from or result in different modes of cognitive processing. More important (for the

language user) may be the differences in the distribution of the two elements.

We are not suggesting that *into* and *enter*, and spatial prepositions and motion/location verbs more generally, are semantically identical. Rather, we are claiming that Cognitive Grammar can account for their different categorial status even if we ignore summary and sequential scanning. Specifically, on the reasonable assumption that *into* (in its motion senses, compare fn.4) tends to co-occur with a motion (or transfer) verb such as *go*, *come*, *walk*, or *put*, the usage-based model leads one to expect that the trajector of *into* is not simply an entity that moves into a container (as is claimed by Langacker, see section 2.1), but is actually elaborated by an entity involved in a process of motion symbolised by the verb (see also Broccias 2003a: 279–280, 2003b). Consider Figure 8, below, which represents the semantic pole of the sentence *She walked into the cinema* (the notational conventions are those used by Broccias 2003a for so-called change constructions). Figure 8b is a compacted version of 8a.

Figure 8

The sentence at hand is analysed as involving the conceptual merger of two components, see Figure 8a: one symbolised by the subject-verb string *She walked* (depicted as the lower box in Figure 8a), and the other symbolised by the prepositional phrase *into the cinema* (the upper box). Note that the dashed line establishes a correspondence between the trajector of the preposition *into* on the one hand, and the subject of the verb, *She*, and derivatively the whole event *She walked*, on the other. Now, given that language users arrive at their semantic analysis of dynamic prepositions such as *into* on the basis of utterances like the one at hand, i.e. utterances where the prepositional phrase depends on some motion verb, it seems reasonable to suggest that, as a generalisation over all these usage events, the trajector of these prepositions is put in correspondence with the trajector of some (schematic) motion event. In other words, a (unspecified) motion event forms part of the base of *into* and similar prepositions. This clearly distinguishes it semantically from *enter* and other motion verbs, which themselves foreground the motion event, and thus do not feature an “extra” component as part of their meaning. Figures 9a and 9b below illustrate our analysis of *into* vs. *enter*. The event symbolised by the squiggly arrow is not in bold in Figure 9a (as

opposed to Figure 8a) because the process it intends to represent is not profiled by *into* but, rather, is part of its base.²¹

Figure 9

It should be clear that while this analysis does not rely on the distinction between summary and sequential scanning, it is nonetheless entirely in the spirit of Cognitive Grammar as a usage-based model: the difference in semantics between *into* and *enter*, and between dynamic prepositions and motion verbs more generally, emerges in the speaker's grammar as a result of differences in usage.²²

Some support for the analysis of *into* as including reference to a schematic motion event is provided by certain "elliptical" patterns in non-standard varieties of English. Certain regional varieties of English may rely on the preposition alone (not necessarily *into*) to code motion, leaving the verb unexpressed, see examples (17a–b), the former taken from Preston (2005), who comments on some (unspecified) dialect of American English, the latter from the BNC. The suggestion that aspects of semantic structure need not always be overtly expressed if they are already schematically present in the utterance, and can be filled in using the linguistic or situational context, is the accepted Cognitive Grammar perspective on what is traditionally known as ellipsis (see e.g. the account of *Jerry will* vs. *Jerry will complain* in Langacker 1991: 491–492).

- (17) a. I need in the house.
 b. "And you want into his knickers," he added a little laugh to put Gerry at ease. (BNC BN1 1071)

The idiomatic phrases in (18) illustrate the same idea, but in these cases the semi-auxiliaries are absent as well:

- (18) a. Off to bed!
 b. Into the fray!²³

An interesting question is, of course, how the lexemes *enter* and *into* arose in the first place, given that they profile the same dynamic scenario.²⁴ The issue could be generalised to similar cases such as *cross* and *across*. Our view would simply be that they perform different functions. *Into* allows speakers to talk about events of entering while at the same time specifying manner of

motion (e.g. *She walked into the cinema*). *Enter*, by contrast, may be used to refer to these events without specifying manner.

6. Conclusion

The discussion has highlighted the problematic nature of summary and sequential scanning, whose existence seems to be needed primarily in order to achieve internal coherence within a lexicalist, semantics-driven linguistic theory. For example, the discussion of infinitival complement patterns with causative verbs may not gain any obvious advantage from the inclusion of summary vs. sequential scanning as an additional dimension of variation. In fact, it would raise questions that are very difficult to answer. The combination of (an improved version of) syntactic/semantic binding and token frequency seems to provide the necessary and sufficient ingredients for a satisfactory analysis. Similarly, *enter* and *into* can be distinguished without recourse to summary and sequential scanning. In the spirit of the usage-based model, the context of use of the two lexemes in question is incorporated into their schematic representations, thus allowing us to account for their different distributions notwithstanding their perceived semantic similarity.

Cognitive Grammar, although being a lexicalist, semantics-driven model as was shown in section 2, indeed recognises the importance of distributional facts for the emergence of grammar. The case of *into* vs. *enter* shows that the view of grammar as a semantics-driven model and the view of grammar as a usage-based model (or corpus model, in the words of Taylor 2003a, b) may go hand in hand, but note that on our analysis the lexical semantics are clearly grounded in the constructions in which the words in question occur. Note, also, that the various infinitival patterns observed for causatives point to the difficulty of always parcelling meaning into minimal discrete components (so that for example we can say that all *to*-infinitives involve summary scanning and all bare infinitives sequential scanning). Taking a lexicalist perspective at the expense of attention to the usage data is potentially harmful because it may confuse the professional linguist with the language user and result in a theory where internal coherence is no longer guaranteed. On balance, then, the lexicalist aspect of the nature of Cognitive Grammar (where the dictum “every form has a meaning” holds sway) compares unfavourably with more radically constructional models like Croft’s (2001) or Gries and Stefanowitsch’s (e.g. Gries and Stefanowitsch 2004;

Stefanowitsch and Gries 2003). Both models recognise a continuum between lexical and constructional meaning (mediated for example by the strength with which individual lexemes are “attracted to”, i.e. occur in, constructions). Furthermore, if the proposed meaning of a form escapes experimental validation and lacks empirical relevance, then it is probably better not to regard it as a cognitively real representation — useful though it might be for the professional linguist’s theory-internal purposes. It is also conceivable that some structures cannot (any longer) be assigned a well-defined meaning on their own but are needed simply to guarantee the conceptual integrity of a construction. For example, it is unlikely that speakers have a representation of an overarching category “subject” along the lines of Langacker (1987), i.e. as a primary figure (see also Croft 2001 on the lack of a single schematic, i.e. globally valid, characterisation of subject; see also Broccias 2006). In terms of the linguistic data we have looked at in this paper, to the extent that the loss of infinitival *to* in some complements was (partly) the result of high token frequency, that seems to suggest that the infinitive marker *to* was (re-)analysed in such a way that its link with a discrete bit of meaning (presumably, a low degree of binding, defined in terms of the parameters outlined in section 3, above) was to some extent lost. (For a careful discussion of the sense in which constructions may be single Gestalts as opposed to being transparent in a perfect morpheme-by-morpheme way see also the discussion in Croft and Cruse 2004: 249–254 of Nunberg *et al.* 1994). In sum, it may be the case that abstract representations are sometimes just professional linguists’ constructs. Our position in this connection is thus similar to Croft’s suggestion in the monosemy vs. polysemy debate, i.e. that “[s]peakers do not necessarily make the relevant generalizations, even if clever linguists can” (Croft 1998: 168).

Another important conclusion stemming from the problematic nature of summary and sequential scanning is that, if we do not include them into the framework, some Cognitive Grammar analyses, e.g. of auxiliaries, adverbs (see Nakamura 1997) and light verbs (see the contrast between *take a fall* vs. *fall* mentioned in Langacker 1987: 146), need rethinking since they are to some extent based on these notions. More in general, the dual nature of Cognitive Grammar — the contrast between the theory-driven requirement of full semantic analysability and the role of entrenchment in the representation of linguistic structures in the speaker’s mind — should be explored in more depth by Cognitive

Grammar practitioners. This problem, we believe, lies at the heart of the construction of a psychologically plausible, and hence truly cognitive, grammar.

We would like to conclude by observing that Langacker (2006, p.c.) suggests that the summary and sequential modes of scanning should actually not be seen as two mutually exclusive cognitive processes, but rather as opposites on a *continuum*. Furthermore, he argues that while at the highest level of interpretation of a clause (i.e. the level of the matrix verb), the difference between the modes may be relatively significant, on lower (embedded) levels the modes play an increasingly less prominent role (see also Langacker 1991: 440–441). The view of scanning modes as non-discrete and as not independent from the clausal context is to some extent closer to the more emphatically usage-based perspective we have argued for in this paper, and has far-reaching implications for all the issues we raised.

In relation to complementation, which we illustrated mainly with reference to causatives (section 3), Langacker's (2006, p.c.) suggestion means that bare vs. *to*-infinitival strategies do not necessarily invoke different representations of the lower clause event at all: the scanning mode is pretty much determined by the matrix verb. This would nullify our objections to different construals for different complementation modes (synchronically and diachronically)—although it would also leave this structural contrast between complements unaccounted for.

However, when we turn to the distinction between dynamic prepositions such as *into* and motion verbs such as *enter* things seem less clear. If the scanning mode of a clause is essentially determined by the (matrix) verb—which is almost always finite—then the question arises as to how, in usage, dynamic prepositions could come to be associated with summary scanning. To see that this is so, consider that a preposition such as *into* will usually co-occur with a motion verb, as in *She walked into the cinema* (see section 5). Now, if the motion verb, in this case *walked*, effects sequential scanning of the situation portrayed by the clause, then it is hard to see whence *into* would derive its summary scanning meaning. One might wish to argue that *into* in isolation features summary scanning, which is overridden when combined with a finite verb, but we submit that *into* is not normally encountered by the language user in isolation (see (18b) for an exception), and that therefore an analysis along those lines would be misguided.

To sum up, the view of scanning as gradient — both in terms of the summary vs. sequential difference and in terms of decreasing prominence from higher to lower levels of clausal organisation — may be an important step towards solving some of the problems associated with a more crude interpretation of the distinction between the two proposed modes, but it remains unclear whether Cognitive Grammar, as a usage-based model, is well served by incorporating the distinction at all.

Received 19 December 2005

Revision received 8 January 2007

Università di Genova
Lancaster University

Notes

- * We would like to thank Bill Croft, Ewa Dabrowska, René Dirven, Adele Goldberg, Stefan Gries, Ronald Langacker and two anonymous referees for their valuable and detailed comments on various versions of this paper. Thanks also go to Catherine Harris, Thomas Egan, Zeki Hamawand and Teenie Matlock for sending us and/or commenting on some of their recent work. All errors, of course, are our own.
1. Langacker (1987: 251) claims that in the case of sequential scanning processing time corresponds to conceived time. In summary scanning this is not the case.
 2. Langacker (1987: 251) suggests that the build-up phase may have no great cognitive salience. Nor is it necessarily the case “that any substantial amount of processing time is required. The span [...] may well be instantaneous for most practical purposes, and it may be construed as a single point in processing time with respect to higher levels of organization” (*ibid.*). However, anticipating the argument in the main text, we note that it is difficult to see at the moment how such claims can be confirmed empirically.
 3. Other prepositions that are analysed as profiling complex atemporal relations include across in sentences such as *Harvey crawled across the table* (Langacker 1991: 217). However, according to Langacker (1991: 217–218), *across* is polysemous in that it can also profile a simple (stative) relation, i.e. in cases where there is no objective motion, e.g. *A famous movie star is sitting across the table* (1991: 217).
 4. It should be borne in mind that we are discussing concrete (i.e. spatially dynamic) uses of the preposition *into*. Non-motion, non-literal examples like *She is into Construction Grammar* do not therefore bear on our discussion.
 5. René Dirven (2005, p.c.) sharply observes that Langacker does not always subsume all to-infinitives under this summary scanning characterisation: “**in certain of its uses**, including its role in forming noun modifiers (e.g. *She was the first person to enter the cave*) [*to*] derives from the processual *enter* a predication that, like nonfinite noun modifiers in general, is both relational and atemporal; *to enter* is thus adjectival rather than verbal, though it clearly profiles all the component states of its base process” (Langacker 1987: 249, emphasis added). The emphasised phrase suggests that some *to*-infinitives

- may not involve summary scanning. Interestingly, in the next subsection he refers to “complex atemporal relations (e.g. *to enter*)”, thereby apparently choosing not to make this qualification.
6. Of course, one could argue that Langacker’s analysis just offers a *schematic* description of the three types of non-finite verb forms and that the purposive meaning associated with the *to*-infinitive might be the prototypical instantiation of the *to*-infinitive schema. But we will see below that there are other reasons to doubt the relevance (and even the existence) of the crucial ingredient for its schematic characterisation, namely summary scanning.
 7. Duffley goes on to suggest that the connection Langacker (1991: 91) makes between summary scanning and the ability to function as a noun modifier leaves unexplained the ungrammaticality of sentences where the bare infinitive is used in this function, e.g. **The only person be named was Theresa*. However, as should be clear from the quotation from Langacker (2002) above, it is actually *to*-infinitives not bare stems that are analysed in terms of summary scanning. And *to*-infinitives can indeed function as noun modifiers, cf. *The only person to be named was Theresa*. Finally, it must be pointed out that Duffley does not offer any alternative account of the phenomena discussed under the rubric of summary and sequential scanning.
 8. It should also be observed that Hamawand uses the terms *temporal* and *atemporal* in a different way from Langacker, i.e. in a more traditional fashion. Hamawand, in connection with *for-to* complements for example, claims that they convert “a *temporal* relation [e.g. what is expressed as a *that*-clause] into an *atemporal* one [...]”. Atemporal denotes a complement clause that is not grounded in time” (Hamawand 2003a: 176; emphasis in the original). According to Hamawand’s definition, both *to*-infinitives and bare infinitives should be classified as atemporal since neither is obviously grounded in time. However, in Langacker’s terminology, *to*-infinitives are said to be atemporal (in that they involve summary scanning) and bare infinitives temporal (in that they involve sequential scanning).
 9. BNC stands for *British National Corpus*, a 100 million word corpus of spoken and written Present-day English; for more information see e.g. Aston and Burnard (1998). The sequence of letters and numbers following this and subsequent examples identifies the location of the relevant sentence in the corpus.
 10. We are grateful to Nick Smith for drawing our attention to Charleston (1960).
 11. We would like to thank one of the two anonymous reviewers for raising this point.
 12. Two more examples of construction variation are seen in the complementation of the verb *help*, as in *Sally helped him (to) do the washing-up*, and of perception verbs, as in *I saw them (to) be obnoxious* (Bolinger 1974: 66–67). Another instance of variation involves the selection of an infinitive vs. a present participle, e.g. *We saw the ship sink/sinking* (Langacker 1991: 442). As Langacker himself points out (1991: 443) the crucial difference between the two variants resides not in scanning but in the perspective taken on the lower clause event, i.e. whether it is external/global, hence requiring an infinitive, or internal/local, hence requiring a present participle (see also section 2.2, above).
 13. As yet another point of concern, one of the anonymous reviewers and Adele Goldberg (2006, p.c.) suggest that summary and sequential scanning are essential in Langacker’s account of the difference between *each* and *every*. In actual fact, Langacker argues that *each*, unlike *every*, involves the operation

- of “sequential examination” (2003: 6, 2005: 192). Nowhere does he equate sequential examination with sequential or summary scanning.
14. Givón also mentions this factor but he only applies it to non-implicative causatives: “non-implicative verbs can already be ranked according to *Likelihood of manipulator’s authority being challenged by the manipulee*, with ‘tell’ coding less challenge and ‘order’, ‘ask’, ‘demand’ coding more” (1980: 368).
 15. The FLOB corpus is a 1 million word collection of British English written prose compiled in the 1990s. More information can be found at <http://khnt.hit.uib.no/icame/manuals/>.
 16. The Helsinki Corpus (diachronic part) is a 1.5 million word corpus covering the period from c750 to 1710; for an elaborate description the reader is referred to e.g. Kytö (1991).
 17. Harris devised a computer experiment where she asked subjects to hit a key when they had understood the meaning of a sentence. Her dependent measure was “understanding time”. However, this method proved insensitive to substituting different types of verbs which might be scanned summarily or sequentially (under the hypothesis that summary scanning is faster than sequential scanning).
 18. We are grateful to Ewa Dabrowska for raising the related question as to how children succeed in linking relevant linguistic expressions to the appropriate scanning mode. More specifically, assuming that summary and sequential scanning are innate cognitive abilities, it is not clear how a child could activate the correct scanning mode when, for example, the caregiver says *Into the box!*
 19. Literate language production is in some ways different from speech; see e.g. Bereiter et al. (1988), Rosenbaum (1990), Rumelhart and Norman (1982) for writing and typing.
 20. The fact that different levels, such as semantics, syntax, phonology and phonetics are distinguished (contrary to the Cognitive Grammar postulate that only two poles, the semantic pole and the phonological pole, both belonging to conceptual space, are needed) is not pivotal to our argument. The crucial point, as is made clear below, is the bottom-up vs. top-down processing issue.
 21. Of course there also exists an intransitive variant of *enter* (e.g. *This word entered into the English language in the 14th century*) which necessarily incorporates the semantic structure symbolised by *into* as part of its base.
 22. The proposed analysis of *into* vs. *enter* is not intended as a comprehensive characterisation of (dynamic) prepositions on the one hand, and verbs, on the other. For a fuller usage-based account we refer to Croft’s parts-of-speech analysis (1991: Chapter 3, 2001: Chapter 2). Here, the major word classes noun, adjective and verb are analysed both in terms of what they describe—viz. objects, properties and actions, respectively—and in terms of the propositional act they are typically used for—viz. reference, modification and predication, respectively. Adpositions are not explicitly discussed in these semantic-pragmatic terms (see Croft 1991: 144–146), but one may argue that they refer to relations between entities, and that their propositional act function corresponds to what Croft calls “situating” (1991: 111–112), i.e. construing the (literal or metaphorical) position of an entity relative to some background dimension in space or time. One might wish to observe at this point that Croft draws a parallel between his own analysis of verbs and Langacker’s account in terms of sequential scanning (1991: 106–107, 121, 123), thereby endorsing the sequential vs. summary scanning distinction.

However, note first that this parallel only applies to dynamic non-stative verbs (1991: 107). Second, and more important, Croft actually prefers to interpret sequential and summary scanning not as the two Langackerian modes of processing a situation, but instead as terms that may be used to describe the ephemeral nature, in terms of discourse processing, of an event described by a verb, as opposed to the more permanent character of opening and maintaining a cognitive file, which occurs when nouns are processed (Croft 1991: 121). This is clearly rather different from Langacker's perspective.

23. These elliptical patterns are not a novelty in English, see e.g. the following examples from Middle English, taken from Mustanoja (1960: 543, 510, respectively):

- i. þat ever dard To hym (*Pearl*, 609–610)
'that ever dared (to go) to him'
- ii. and took hire leve, and hom ... (Chaucer, *Troilus and Criseyde*, i, 126)
'and took her leave, and (went) home'

In other Germanic languages such as Dutch and German the elliptical construction is probably even the preferred pattern with modals, at least in spoken varieties: Du. *Nu moet ik naar huis*, G. *Jetzt muss ich zu Hause* 'I must go home now'.

24. We are grateful to one of the two anonymous reviewers for pointing this out to us.

References

- Anderson, John R.
1995 *Cognitive Psychology and Its Implications*. New York: W.H. Freeman.
- Aston, Guy and Lou Burnard
1998 *The BNC Handbook: Exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.
- Barsalou, Lawrence W.
1992 *Cognitive Psychology. An Overview for Cognitive Scientists*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bereiter, C., P.J. Burtis and M. Scardamalia.
1988 Cognitive operations in constructing main points in written comprehension. *Journal of Memory and Language* 27, 261–278.
- Bolinger, Dwight
1974 Concept and percept: Two infinitive constructions and their vicissitudes. In *World Papers in Phonetics: Festschrift for Dr. Onishi's Kiju*. Tokyo: Phonetics Society of Japan, 65–91.
- Broccias, Cristiano.
2003a *The English Change Network. Forcing Changes into Schemas*. Berlin: Mouton de Gruyter.
2003b Quantity, causality and temporality in change constructions. In Nowak, Pawel M., Corey Yoquelet, and David Mortensen (eds.), *Proceedings of the Twenty-ninth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetic Sources of Phonological Patterns*. Berkeley, CA: Berkeley Linguistics Society, 67–78.

- 2006 Cognitive approaches to grammar. In Kristiansen, Gitte, Michel Achard, René Dirven, and Francisco J. Ruiz de Mendoza Ibañez (eds.), *Cognitive linguistics: Current applications and future perspectives*. Berlin: Mouton de Gruyter, 81–115.
- Bybee, Joan L. and Joanne Scheibman
1999 The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics* 37, 575–596.
- Charleston, Britta M.
1960 *Studies on the Emotional and Affective Means of Expression in Modern English*. Bern: Francke.
- Clark, Herbert H. and Eve V. Clark
1977 *Psychology and Language*. New York: Harcourt Brace Jovanovich.
- Comrie, Bernard
1976. *Aspect*. Cambridge: Cambridge University Press.
- Cristofaro, Sonia
2003 *Subordination*. Oxford: Oxford University Press.
- Croft, William
1991 *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: Chicago University Press.
1998 Linguistic evidence and mental representations. *Cognitive Linguistics* 9, 151–173.
1999 Some contributions of typology to cognitive linguistics, and vice versa. In Janssen, Theo and Gisela Redeker (eds.), *Cognitive Linguistics: Foundations, Scope, and Methodology*. Berlin: Mouton de Gruyter, 61–93.
2001 *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Croft, William and D. Alan Cruse
2004 *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Dell, Gary S.
1986 A spreading activation theory of retrieval in sentence production. *Psychological Review* 93, 283–321.
- Dixon, R.M.W.
2000 A typology of causatives: form, syntax and meaning. In Dixon, R.M.W. and Alexandra Y. Aikhenvald (eds.), *Changing Valency. Case Studies in Transitivity*. Cambridge: Cambridge University Press, 30–83.
- Duffley, Patrick J.
1992 *The English Infinitive*. New York: Longman.
2003 The gerund and the *to*-infinitive as subject. *Journal of English Linguistics* 31, 324–352.
2006 The English gerund-participle in Cognitive Grammar. *Lacus Forum* XXXII, 301–306.
- Egan, Thomas.
In prep. *Non-finite Complementation: A Usage-Based Study of Infinitive and -ing Complement Clauses in English*. Amsterdam: Rodopi.
- Fillmore, Charles J.
1972 Subjects, speakers, and roles. In Davidson, Donald and Gilbert Harman, (eds.), *Semantics of Natural Language*. Dordrecht: Reidel, 1–24.

- Francis, Elaine
 2000 Two perspectives on the grammar of possession. *Language Sciences* 22, 87–207.
- Fromkin, Victoria A.
 1971 The non-anomalous nature of anomalous utterances. *Language* 47, 27–52.
- Fromkin, Victoria A. (ed.)
 1973 *Speech Errors as Linguistic Evidence*. The Hague: Mouton.
- Garrett, Merrill F.
 1975 The analysis of sentence production. In Bower, Gordon H. (ed.), *The Psychology of Learning and Memory: Advances in Research and Theory*, vol. 9. New York: Academic, 133–177.
 1980 The limits of accommodation. In Fromkin, Victoria A. (ed.), *Errors in Linguistic Performance*. New York: Academic, 263–271.
 1988 Processes in language production. In Newmeyer, Frederick J. (ed.), *Linguistics: The Cambridge Survey*, vol. 3. *Language: Psychological and Biological Aspects*. Cambridge: Cambridge University Press, 69–96.
- Givón, Talmy.
 1980 The binding hierarchy and the typology of complements. *Studies in Language* 4, 333–377.
 1990 *Syntax. A Functional-Typological Introduction*, vol. 2. Amsterdam: John Benjamins.
- Gries, Stefan and Anatol Stefanowitsch
 2004 Covarying collexemes in the *into*-causative. In Achard, Michael and Suzanne Kemmer (eds.), *Language, Culture, and Mind*. Stanford: CSLI Publications, 225–236.
- Haiman, John
 1985 *Natural Syntax: Iconicity and Erosion*. Cambridge: Cambridge University Press.
- Hamawand, Zeki
 2002 *Atemporal Complementisation Clauses in English: A Cognitive Grammar Analysis*. München: Lincom Europa.
 2003a *For-to* complement clauses in English: A Cognitive Grammar analysis. *Studia Linguistica* 57, 171–192.
 2003b The construal of atemporalisation in complement clauses in English. *Annual Review of Cognitive Linguistics* 1, 59–85.
- Hollmann, Willem
 2003 Synchrony and diachrony of English periphrastic causatives: a cognitive perspective. Unpublished doctoral dissertation, University of Manchester, Manchester.
 2005 The iconicity of infinitival complementation in Present-day English causatives. In Maeder, Constantino, Olga Fischer and William J. Herlofsky (eds.), *Outside-in—Inside-out. Iconicity in Language and Literature*, vol. 4. Amsterdam: John Benjamins, 287–306.
- Jackendoff, Ray S.
 1972 *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Just, Marcel A. and Patricia A. Carpenter
 1980 A theory of reading: from eye fixations to comprehension. *Psychological Review* 87, 329–354.

- Kytö, Merja
 1991 *Manual to the Diachronic Part of the Helsinki Corpus of English Texts: Coding Conventions and Lists of Source Texts*. Helsinki: Department of English, University of Helsinki.
- Langacker, Ronald W.
 1987 *Foundations of Cognitive Grammar*, vol. 1. *Theoretical Prerequisites*. Stanford, CA.: Stanford University Press.
 1991 *Foundations of Cognitive Grammar*, vol. 2. *Descriptive Application*. Stanford, CA: Stanford University Press.
 1992 Prepositions as grammatical(izing) elements. *Leuvense Bijdragen* 81, 287–309.
 1999 *Grammar and Conceptualization*. Berlin: Mouton de Gruyter.
 2002 *Concept, Image and Symbol*. 2nd edn. Berlin: Mouton de Gruyter.
 2003 One any. *Korean Linguistics* 18, 65–105. [page reference in the main text to the MS version]
 2005 Dynamicity, fictivity, and scanning: The imaginative basis of logic and linguistic meaning. In Pecher, Diane and Rudolf A. Zwaan (eds.), *Grounding Cognition: The Role of Perception and Action in Memory, Language and Thinking*. Cambridge: Cambridge University Press, 164–197.
- Levelt, W.J.M.
 1989 *Speaking: From Intention to Articulation*. New York: Wiley.
- Matlock, Teenie
 2004a Fictive motion as cognitive simulation. *Memory and Cognition* 32, 1389–1400.
 2004b The conceptual motivation of fictive motion. In Radden, Günter and Klaus-Uwe Panther (eds.), *Studies in Linguistic Motivation*. Berlin: Mouton de Gruyter, 221–248.
- Matlock, Teenie, Michael Ramscar, and Lera Boroditsky
 2005 The experiential link between spatial and temporal language. *Cognitive Science* 29, 655–664.
- Mittwoch, Anita
 1990 On the distribution of bare infinitive complements in English. *Journal of Linguistics* 26, 103–31.
- Mustanoja, Tauno F.
 1960 *A Middle English Syntax*, vol. 1. Helsinki: Société Néophilologique.
- Nakamura, Wataru
 1997 A cognitive approach to English adverbs. *Linguistics* 35, 247–287.
- Nunberg, Geoffrey, Ivan A. Sag and Thomas Wasow
 1994 Idioms. *Language* 70, 491–538.
- Pelyvás, Péter
 2006 Subjectification in (expressions of) epistemic modality and the development of the grounding predication. In: Athanasiadou, Angeliki, Costas Canaki, and Bert Cornillie (eds.), *Subjectification: Various Paths to Subjectivity*. Berlin: Mouton de Gruyter, 121–150.
- Preston, Dennis R.
 2005 How can you learn a language that isn't there? In Dziubalska-Kolaczyk, Katarzyna and Joanna Przedlacka (eds.), *English Pronunciation Models: A Changing Scene*. New York: Peter Lang, 37–58.

- Reddy, D. Raj (ed.)
 1975 *Speech Recognition*. New York: Academic.
 1980 Machine models of speech perception. In Cole, Ronald A. (ed.), *Perception and Production of Fluent Speech*. Hillsdale, NJ: Lawrence Erlbaum Associates, 215–242.
- Rosenbaum, David A.
 1990 *Human Motor Control*. New York: Academic.
- Rumelhart, David E. and Donald A. Norman
 1982 Simulating a skilled typist: a study of skilled cognitive-motor performance. *Cognitive Science* 6, 1–36.
- Sandra, Dominiek and Sally Rice
 1995 Network analyses of prepositional meaning: Mirroring whose mind—the linguist’s or the language user’s? *Cognitive Linguistics* 6, 89–130.
- Stefanowitsch, Anatol and Stefan Gries
 2003 Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8, 209–243.
- Talmy, Leonard
 1976 Semantic causative types. In Shibatani, Masayoshi (ed.), *Syntax and Semantics 6: The Grammar of Causative Constructions*. New York: Academic Press, 43–116.
 1983 How language structures space. In Pick, Herbert and Linda P. Acredolo (eds.), *Spatial Orientation: Theory, Research, and Application*. New York: Plenum Press, 225–282.
 1985 Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*, vol. 3. *Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press, 57–149.
 1988 Force dynamics in language and cognition. *Cognitive Science* 12, 49–100.
 2000 *Toward a Cognitive Semantics*, vol. 1. *Concept Structuring Systems*. Cambridge, MA: MIT Press.
- Taylor, John
 2002 *Cognitive Grammar*. Oxford: Oxford University Press.
 2003a Polysemy’s paradoxes. *Language Sciences* 25, 637–655.
 2003b Cognitive models of polysemy. In Nerlich, Brigitte, Zazie Todd, Vimala Herman, and David Clarke (eds.), *Polysemy. Flexible Patterns of Meaning in Mind and Language*. Berlin: Mouton de Gruyter, 31–47.
- Tomasello, Michael
 2003 *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Verhagen, Arie and Suzanne Kemmer
 1997 Interaction and causation: Causative constructions in modern standard Dutch. *Journal of Pragmatics* 27, 61–82.
- Visser, Fredericus Th.
 1973 *An Historical Syntax of the English Language*. Part Three, Second Half. *Syntactical Units with Two and with More Verbs*. Leiden: E.J. Brill.
- Ungerer, Friedrich and Hans-Jörg Schmid
 1996 *An Introduction to Cognitive Linguistics*. London: Longman.

Wierzbicka, Anna

1975 Why "kill" does not mean "cause to die": The semantics of action sentences. *Foundations of Language* 13, 491–528.