

# ASPECT, ASPECTUAL CLASS, AND THE TEMPORAL STRUCTURE OF NARRATIVE

Alexander Nakhimovsky

Department of Computer Science  
Colgate University  
Hamilton, NY 13346 USA

This paper consists of two parts. The first part discusses commonsense knowledge about events as manifested in language. Three kinds of knowledge are identified: compositional, durational, and aspectual. Compositional knowledge concerns internal structuring of events into preparatory, initial, main (the body), final, and resulting stages. Durational knowledge concerns durational relations between events and stages of the same event. Durational knowledge can be expressed as qualitative dependencies among the parameters of the event and as its time scale. The notion of time scale is introduced and related to shared cyclical events (time units).

In discussing aspectual knowledge three notions are distinguished: aspect as a grammatical category of the verb, implemented by affixes, auxiliaries, and such; aspectual class, which is a characteristic of a lexical meaning; and the aspectual perspective of the sentence determined by the position of the Reference Time (RT) with respect to the event described by a finite clause. I argue that an aspectual classification of situations evolving in time should be based on such considerations as the kinds of resources they consume and the goals they achieve. A detailed classification of instantaneous and noninstantaneous events is developed.

The second part of the paper discusses how this knowledge is employed in understanding extended narratives. Temporal discontinuities, in conjunction with other kinds of discontinuities identified in the paper, signal boundaries between discourse segments; within each segment, all three varieties of temporal knowledge help establish the temporal relations among narrated events.

## 1 INTRODUCTION

This paper consists of two parts.<sup>1</sup> The first part (sections 2–5) discusses aspect and other varieties of commonsense knowledge about time as manifested in language. The second part (sections 6–8) discusses how this knowledge is employed in understanding extended narratives. This preliminary section spells out the paper's principles and aspirations.

A major influence on this paper has been the study of commonsense knowledge and reasoning in the paradigm established in Hayes 1985 (1979), Hobbs and Moore 1985, and Hobbs et al., 1985, 1986. The paper proceeds from the premise that commonsense studies and natural language semantics are closely interdependent: natural language semantics is essentially incomplete without a theory of the commonsense world, for which natural language provides key empirical data. These data are of two kinds, lexical and grammatical. So far, mostly

lexical data have been used in work on knowledge representation: the contexts of a given word, the structure of semantic fields, and metaphorical uses (Hobbs 1984:284, 1985a:3). This paper suggests that in working on core knowledge (time, space, causality), it is imperative to recognize grammatical categories as powerful, and language specific, tools of organizing conceptual space.

Among grammatical categories particularly important are those that allow, and indeed force, the language user to choose between alternative ways of viewing the same object or situation. I will call such categories *subjective*. A paradigmatic example of a subjective grammatical category is aspect. (In fact, the term *aspect* came into being as an attempt to translate the Russian term *vid* ("view"); see Lyons 1977:705.) Subjective grammatical categories are important for two reasons. First, they suggest restrictions on the formalism: since

Copyright 1988 by the Association for Computational Linguistics. Permission to copy without fee all or part of this material is granted provided that the copies are not made for direct commercial advantage and the *CL* reference and this copyright notice are included on the first page. To copy otherwise, or to republish, requires a fee and/or specific permission.

0362-613X/88/010029-43\$03.00

language users can easily switch between alternative views of the same entity, the formalism should provide for alternative representations related by a *simple and general operation*. (Bunt 1985:39–40 makes a similar point with respect to the opposition mass-count.) Second, subjective grammatical categories show that the corresponding conceptual distinctions reflect our alternative ways of viewing things rather than intrinsic properties of things. There are intrinsic properties that interact with subjective distinctions, and in a language that does not have a well-articulated subjective grammatical opposition, the intrinsic and the subjective become easily confused. An example of such a confusion is, I believe, the well-known taxonomy of events, processes, and states (Mourelatos 1981), which is discussed in Section 4.3.

It follows that in designing the primitives of our semantic representations it is essential to look at data across languages, paying particular attention to subjective grammatical categories. (They are easy to identify because they are notoriously difficult to learn how to use correctly, possibly because of having to internalize a new global operation on the mental model.) Decades of linguistic research can thus be brought to bear on the enterprise of formalizing commonsense knowledge. Conversely, this enterprise makes it possible to pose with greater precision the old problem of how language, and in particular its grammatical categories, influences our thought and vision of the world.

Formalizations of the commonsense may or may not turn out to be language and culture specific, but they will have to have features particular to our planet and the species. Such accidental parameters of human existence as the duration of a year, the length of an arm, and the temperature of the human body are universal constants that play an important role in our reasoning: they determine how we factor a continuous parameter like temperature or time into a quantity space (in the sense of Forbus 1985:80–1). Therefore, and contrary to a widespread view, any rigorous semantics of natural language has to concern itself with definitions of extensional concepts. Such definitions obviously belong in the lexicon. In this paper, lexicon is understood to be a repository of both grammatical and commonsense knowledge, indexed by lexical items. This view of the lexicon, advocated also in Hobbs et al., 1985, has important antecedents in Soviet and European linguistics (Apresjan 1974, Wierzbicka 1985).

A second major influence on this paper is the notion of **computational modeling** of natural-language semantics (Hendrix and Moore 1979), i.e., the idea that semantical objects associated with natural-language expressions are (typed) data structures, together with inference procedures that operate on them. These data structures form an internal representation language that may or may not in turn be endowed with a semantics of its own. In what follows I assume that semantical objects are frames and collections of frames, and the

basic types include **frame** and **slot label** (cf. Creary and Pollard 1985, and Hirst 1987, ch.3). There are sortal predicates, or simply sorts, that are defined on frames and organize them into an ISA hierarchy.

The first cut in the hierarchy is between frames representing objects and frames representing situations persisting or evolving in time. (Loosely following Hayes 1985, I call the latter **histories**.) Orthogonal to this distinction is the one between generic frames or types, and their instances or tokens: tokens, but not types, have a slot for a spatio-temporal location. I will thus speak of **history types (h-types)** and **history tokens (h-tokens)**; when the type-token distinction is irrelevant I will simply say “history.” H-types can be of various degrees of generality (cf. the representations of *read*, *read a book*, *read a book for an hour*; all of these, however, are types of which the representation of *Jon read a book for an hour* is a token). Associated with each h-type is a set of slot labels and selectional restrictions on them, usually a conjunction of sortal predicates (Hobbs et al., 1986).

Whether types or tokens, the semantical objects of natural-language semantics are only partially specified, and “more partial” objects are embeddable, with some provisions for nonmonotonicity, into less partial ones. I assume that representations of lexical items in the lexicon form a discrete level of partiality. The subject of the first part of the paper can then be stated as follows: What is the knowledge about events that is part of lexical meanings? How can it be coded in the h-types of the lexicon, and how is it used as lexical meanings are combined and embedded into more precisely specified structures? In the end, what it all boils down to is an extended argument for some slots in the lexical representations of histories and a few sortal predicates on h-types representing aspectual classes.

Finally, with respect to narrative understanding, a major influence has been the idea that such understanding means *constructing*, or “building a good structure” (Bruce 1981:283), so that the semantic value of a linguistic unit (phrase, sentence) is its contribution to the emerging representation of discourse. This idea, in one of those parallel developments that do not seem to be purely coincidental, emerged at about the same time in the Reader-Response school of literary criticism (e.g., Tompkins 1980), in model-theoretic semantics (Kamϩ 1981, Heim 1982) and in AI approaches to discourse; in combination with the idea of computational modeling it first clearly appears in Webber 1979 (see also Webber, this volume).

## 2 VARIETIES OF TEMPORAL KNOWLEDGE AND COMPOSITION OF H-TYPES

This and the next two sections summarize and elaborate on the proposals of Nakhimovsky 1986, 1987a. Commonsense knowledge about time falls into three categories: compositional, durational, and aspectual. Compo-

sitionally, h-type representations in the lexicon consist of the following components:

- a. Preparatory stage.
- b. Initial stage.
- c. Body.
- d. Final stage.
- e. Resulting stage.

In Nakhimovsky 1986 only components **b–e** were recognized, but Moens and Steedman 1987 (also this volume) convincingly shows that **a** needs to be added, at least when the h-token is a planned human action. Components **b–e** are probably uncontroversial, except in certain cases, discussed below, where they are collapsed together. For example, consider the verb *cross*, as in *cross the road*. It describes movement from one side of a ribbonal surface to another, following a trajectory that is usually at approximately the right angle to the edge of the surface. The stages are: be located on one side, cross the edge, move across, cross the other edge, be located on the other side. So, as Steedman and Moens observe, the clause *When the children crossed the road . . .* can be followed by any of the following:

#### Example 1

- a. They waited for the teacher to give a signal.
- b. They stepped onto its concrete surface as if it were about to open and swallow them.
- c. They were nearly hit by a car.
- d. They reached the other side stricken with fear.
- e. They found themselves surrounded by strangers.

Moens and Steedman argue, correctly I think, that the different temporal relations between the *when* clause and the continuations do not result from the ambiguity, or polysemy, of the temporal connective: the *when* clause simply makes both the entire h-token and all its component stages available for focusing. What Example 1 does show is the weak aspectual typing of English verb forms: the same simple past form *crossed* makes all the stages of the h-type accessible. The most that can be said is that before some, but not all, of the continuations in Example 1, the simple past can be replaced by the progressive. In languages with strong aspectual typing an imperfective verb (form) would *have* to be used in the *when* clause before (1b) or (1c).

This example points up what I believe to be the main function of aspect as a grammatical category: making either an h-type in its entirety, or (some subsequence of) its internal stages, available for focusing. In order to make the notion of focusing more precise I will use the notion of **Reference Time (RT)**, adopted from Reichenbach 1947 but reinterpreted pragmatically: RT is to be thought of as a point or a very short interval that indicates where the current focus of attention is located. It may be obvious that after (1a) RT is located in the middle of the preparatory stage, after (1b) in the beginning of the body right after the initial stage, after (1c) in

the middle of the body, after (1d) in the resulting stage right after the final stage, and after (1e) in the resulting stage. If, by the end of the sentence, the RT is positioned *inside* the body of the h-token described by the sentence I will say that the sentence has the **imperfective aspectual perspective**; if the RT is positioned after the final stage or, more rarely, before the initial stage, I will say that the sentence has the **perfective aspectual perspective**. In Example 1, the initial *when* clause with a verb in the simple past leaves the aspectual sentence perspective underdetermined, which is what I meant by weak aspectual typing of English.

Note that the RT never ends up in the middle of the initial or final stages, because they are too short, much shorter than the body; they belong to a different time scale. This does not mean that they are necessarily instantaneous in an absolute sense: they are simply shorter than the body by at least a constant factor, and in a different sentence they can still be presented imperfectively. Consider Example 2:

#### Example 2

- a. We spent the whole of last week working on Bob's project.
- b1. In the beginning . . .
- b2. As we began working on it . . .
- b3. When we were just beginning our work . . .

The verb *begin* followed by an infinitive or gerund is a standard way of referring to the initial stage of a history. (Sometimes a lexical item is available as well or instead: “to begin to be asleep” is idiomatically expressed as *fall asleep*.) The second sentence of Example 2 thus puts the RT in the middle of the initial stage of the h-token described by the first sentence. Note that if in (2a) we replace the word *week* with *month*, *year*, *decade*, or *night*, the duration of the beginning varies with the duration of the whole while remaining much shorter than the body.

The fact that we can so “zoom in” on different stages of a history has led many to conclude that there is no absolute sense in which events can be described as instantaneous. As Allen and Kautz 1985:253 puts it, “we can always ‘increase the magnification’ and find more structure” (see also Dowty 1986, Kamp 1979). I believe that instantaneousness is an absolute quality determined by our biology: instantaneous events are those that are not perceived by humans as possessing internal structure. Languages select such events for special treatment by disallowing the “imperfective description” of them: one cannot use the imperfective aspect to place the RT in the middle of an instantaneous event, so that *The light was flashing* does not place the RT inside an individual flash. Nor can the *begin* plus gerund construction refer to the initial stage of an instantaneous event: compare *He began crossing the street* with *He began blinking*. The initial stage, the body and the final stage of an individual blink or flash or

any other instantaneous event are thus inaccessible for future focusing, and it makes sense to represent them as all collapsed together.<sup>2</sup>

Note that the class of instantaneous events can be defined from three different perspectives corresponding to the three varieties of temporal knowledge postulated in the beginning of this section. Durationally, instantaneous events belong to a time scale that is below a certain threshold of human perception of time (cf. Fraisse 1984:29–30). Compositionally, they have impoverished internal structure. Aspectually, this shows up in imperfective forms of verbs describing instantaneous events, which are interpreted differently than such forms of verbs describing noninstantaneous histories.

### 3 KNOWLEDGE ABOUT DURATIONS

Durations can be entered in the lexicon in the following three ways that are not mutually exclusive:

- a. For some h-types (e.g., **lecture**, **shower**, **lunch**) we know the duration of their h-tokens to be stable, albeit approximate. Such durations can be entered in the lexicon directly as a fuzzy number, e.g., **lecture** [1 2 hour] (cf. Allen and Kautz 1985). This is a simple and least interesting case.
- b. For most aspectual classes of histories, duration is constrained by qualitative functional dependencies (Forbus 1985) among the participants of the situation; so, the time it takes to read a text depends on its length and genre, and the proficiency of the reader. This case is discussed in Section 4, after the notion of aspectual class is introduced.
- c. For a majority of h-types, constraints on their durations can be expressed using the notion of time scale, introduced below.

In order to measure durations one needs cyclical events. These may be celestial, biological, or man-made. (See Campbell 1986 for a lively discussion.) Time-measurement units provided by cyclical events form several interlocking, multitiered hierarchies. It is commonly assumed in AI literature that the choice of a measurement unit is arbitrary and independent of the interval to be measured, because we can always switch from one unit to another by using axioms of the sort hour = 3600 seconds. Mathematically, of course, this is true, but I don't think language quite works this way: when we say *It took Bob an hour to repair the faucet*, we don't mean that it took him 3600 seconds. Given an interval, we choose a unit to measure it with so that the duration is expressed by a "reasonable" number, that is, not too big. Small fractions are also avoided: we say half a month or a quarter of an hour, but not one-fifth, -sixth, and so on. (Note that a quarter is the smallest fraction for which there is a special word; the names for the rest of them are formed by a regular morphological process.) Under these (psychologically) very under-

standable restrictions it is rarely the case that an interval fit precisely into the measurement unit we use. This results in a certain degree of fuzziness: when we say that something happened a year ago, we typically don't claim to be precise to the day. The same imprecision is observed when we use man-made time units such as hours, minutes, and seconds. The sentences *Joe slept for three hours*, or *Joe slept from 2 to 5* do not claim precision to the minute; in fact, the second sentence can be continued with *He woke up at a quarter after 5 feeling greatly refreshed*, and the narrative does not become self-contradictory.

These observations suggest that a duration expressed by a number in hours becomes expressed by a fuzzy number (an interval) in minutes: [3 3 hour] becomes something like [160 200 minute]. (In a richer theory, the degree of fuzziness would depend on what we know about the source of information.) Since the boundaries of a fuzzy interval are themselves fuzzy, it might seem that we ought to throw in some seconds at the ends of the [250 300] minutes interval. However, the duration would then be expressed in thousands of seconds, and commonsense reasoning is too qualitative to deal with such big numbers. The length of the interval suggests a measurement unit, which, in turn, establishes a certain scale, and durations that are much smaller than this scale are disregarded.

It thus makes sense to talk of units of measurement that are "natural" for a given interval: measured in a natural unit, the length of the interval will not be a very small fraction (greater than some constant *R*) or a very big number (less than some constant *N*). I wouldn't want to commit myself to specific values for *R* and *N*, but psychologists seem to have a fairly clear idea of what they are (see, e.g., Fraisse 1984). The **time scale** of an interval can then be characterized by the sequence of its natural units, a subsequence of all the duration units. An interval's duration can be represented by fuzzy numbers, e.g., [1 1 year] "a year, about a year"; [1.5 1.5 week] "a week and a half"; [2 3 day] "two to three days" and so forth. There are rules that transform fuzzy durations in one unit into fuzzy durations in a smaller unit, but only as long as both units are natural to the interval being measured: [1.5 1.5 week] becomes [9 12 day], but not [200 260 hour].

To capture the intuition that hours are simply irrelevant when we talk about weeks, or days when we talk about years, I introduce the relation **much-greater-than**, which is readily defined in terms of natural units: *i* is much greater than *j* if and only if (iff) the largest natural unit of *j* is smaller than the smallest natural unit for *i*. There are thus two ordering relations defined on durations measured in the same scale: **greater-than** and **much-greater-than**; together, they give rise to an **equi** (indistinguishability) relation (cf. Hobbs 1985, 1986b) defined as follows: two intervals' durations are **equi** iff the smaller of the two durations is much greater than the difference between them. Note that in this scheme of

things the **equi** relation is derived, rather than primitive as in Hobbs's work, and it is not incompatible with **greater-than**. I find it both convenient and intuitively correct that we can treat two intervals as equivalent without losing the knowledge, if we have it, that one is greater than the other.

Note that the notion of natural units presented here is related to McDermott's (1982) *lifetimes*: saying that the natural unit for property **p** is *year* means that the property lasts (or persists) for years. It is also related to Hobbs's (1985a) *grain*. What I propose is that lifetimes and grains are monotonically related: the greater the lifetime, the greater the grain. Thus the same parameter, natural unit, can capture both.

Duration units cut up a continuous parameter (time) into a finite set of quantities. The set of duration units, with the **greater-than** and **much-greater-than** relations on it, thus forms a quantity space for time, in the sense of Forbus 1985:96–7. We can communicate because our reasoning systems employ the same temporal quantity space, whose elements are arbitrary in precisely the same sense in which linguistic signs are arbitrary: each one of them could have been different, yet none of them can be changed by the individual speaker/reasoner, who acquires them in the process of cognitive and linguistic development. The same progression of ideas applies, with some modifications, to other one-dimensional scalars: distance, temperature, price, height/tallness, and speed.

Given the role of time-measurement units in the proposed framework, the task of axiomatizing the clock and the calendar, mentioned in Hobbs 1985, becomes highly important. It involves identifying the time-measurement units, defining them in terms of cyclical events, and establishing the **much-greater-than** relation on them. There is some discussion of the task in Nakhimovsky 1986; here I can give just one illustration of the relevant empirical arguments. Apart from measurement units with single word names, there are multiples of other units, described by phrases, that have themselves become measurement units. *Five minutes* provides an example: we talk about 10 or 15 or 20 minutes much more commonly than 11 or 14 or 17 minutes, and when asked to estimate a duration that is less than an hour we typically use a multiple of five minutes. (Note that this is not true about hours or days.) If we don't, an unusual precision is implied; compare:

### Example 3

- a. I'll be back in 15 minutes.
- b. I'll be back in 14 minutes.

There's much more of a broken promise if I'm one minute late after (3b) than (3a), which indicates a finer grain (smaller scale) in (3b).

Clearly, five minutes have become a unit because a minute is a unit, and we have five fingers on each hand. This is yet another example of the anthropocentrism pervading language and commonsense reasoning, the

same antropo-centrism that was mentioned earlier in connection with instantaneous events.

## 4 ASPECT

### 4.1 GRAMMATICAL ASPECT

In what follows it is essential to keep the following three concepts apart: **aspect** as a grammatical category of the verb, implemented by affixes, auxiliaries, and such; **aspectual class**, which is a characteristic of an h-type or lexical meaning; and the **aspectual perspective of the sentence** determined by the position of the RT with respect to the h-token described by a finite clause. Thus grammatical aspect is associated with linguistic expressions, aspectual class with h-types, and aspectual perspective with h-tokens. It is the aspectual perspective of the sentence that is ultimately important for understanding; both grammatical aspect and aspectual class sometimes uniquely determine, sometimes just strongly constrain, the aspectual perspective. In English, for instance, the progressive aspect guarantees that the sentence perspective is imperfective, except for sentences describing instantaneous events, which, in any language, cannot be presented imperfectively. (As the discussion in the end of Section 2 put it, *He was blinking* or *He started blinking* does not put the RT in the middle of an individual blink.) This does not mean that verbs like *blink* are grammatically perfective.

There are languages, most notably Slavic, where the difference in sentence perspective is hard-wired into verb morphology: simplifying very slightly, every Russian verb is either perfective or imperfective (with the diagnostics being that only imperfective infinitives can follow verbs like *begin* or *continue*), and the morphological feature of the verb determines the aspectual perspective of the sentence. In English, by contrast, simple past is fairly noncommittal as to the aspectual perspective of the sentence, as we saw in the discussion of Example 1. Another way to describe this contrast is to say that Russian has a perfective-imperfective aspectual system, while English has an imperfective-unmarked aspectual system; most aspectual systems seem to belong to one of these two kinds, with about three-to-one advantage for the perfective-imperfective systems in Dahl's (1985) analysis of 64 languages.

A fairly common misconception is to assume that since English has an imperfective, it ought to have a perfective also. Obviously, simple past, without any further qualifications, cannot be considered perfective, in view of such obvious examples as *Bob owned a house on Cape Cod*. However, sentences like *Vanessa read a magazine article about Mongolia* do suggest the perfective perspective: by the end of the sentence the article seems to have been read (cf. Hinrichs 1986:68; Dowty 1986:46–8 for similar examples and discussion). I argue that it would be incorrect to call the verb or the verb phrase in such sentences perfective, because the perfective sentence perspective emerges not so much from

the linguistic properties of the verb phrase as from commonsense knowledge about the situation it describes: a short process that progresses towards a well-defined terminal point beyond which it cannot continue. To see that we are not dealing with a grammatical category but rather commonsense inference, observe that an appropriate context can change the sentence perspective, as in Example 4, where Vanessa's reading is interrupted by a doorbell:

#### Example 4

After supper, Vanessa and Didi sat down in the living room. Vanessa read a magazine article about Mongolia, Didi watched her favorite cartoon on television. Suddenly the doorbell rang.

I conclude, in accord, I believe, with Passonneau (1987 also this volume), that English has no morphological perfective; the perfective effect arises out of the aspectual properties of the history being described, i.e., its aspectual class.

#### 4.2 ASPECTUAL CLASSES OF HISTORIES

An aspectual classification divides lexical meanings according to the internal temporal structuring of the corresponding h-tokens. I believe they should be thought of as intrinsic properties of lexical meanings, independent of context and, in particular, aspectual perspective. Even out of context we know a good deal about how instances of, e.g., reading or sleeping or walking to the store are individuated and how they unfold in time. This knowledge interacts with, but is not determined by, whether the h-token is presented imperfectly, as in *John was walking to the store*.

The major division among noninstantaneous histories, recognized at least since Aristotle, is between process (*energeia*) and state (*stasis*). In recent times, Vendler (1967) proposed a highly influential classification that is still commonly accepted, although the principles of classification have changed. Vendler believed, erroneously, that he was classifying English verbs, rather than their denotations, and he used such language-specific criteria as whether or not a verb has a progressive form (Vendler's statives, such as *know*, don't). In the Taylor-Dowty model-theoretical version (see, e.g., Dowty 1986), the classification is based on the relationship between the truth value of a sentence at an interval and at its subintervals. So, for instance, a sentence S is stative (denotes a state) iff it follows from the truth of S at an interval I that S is true at all subintervals of I (the so-called subinterval property, cf. Dowty 1986:42).

I submit that these criteria cannot possibly capture the real distinctions operative in the workings of human language: these have to relate to something perceived and experienced, rather than truth values (which is not to deny that real distinctions may result in fairly consistent truth-functional properties). It is not accidental that Dowty's own example of a state (*sleep*) contradicts

his definition: we can truthfully say that Bob slept from 10 to 6 even if he got up once to go to the bathroom. My proposal is that we take the physical vocabulary of processes and states seriously, and classify histories according to their internal dynamics, the stability of their parameters, and the resources they consume. (Part of the internal dynamics, in the presence of a conscious agent, is the degree of required attention and volitional control.) We can then note the distinction between states that do not require any resources to sustain themselves (knowing English, owning a house) and those that do (sleep requires a certain amount of sleepiness, which gradually wears out). The subinterval property holds only for zero-resource states and is, in fact, a simple consequence of their other properties: a state that requires no resources, including attention and conscious control, obtains continuously.

An important group of resources are those that are consumed by the unconscious and periodic metabolic processes sustaining a living organism. I call them "generic resources" because they are not specific to any particular history, but rather to all individuals of a given sort. Resource-consuming states all seem to require only generic internal resources, while within processes, there are those that require only generic resources (walking) and those that require process-specific resources as well: reading, for example, requires not only being awake and not too hungry, but also a text to read. This is the first of the two important divisions among processes. The other one separates telic processes from atelic ones.

Telic processes can be defined as processes that have a built-in terminal point that is reached in the normal course of events and beyond which the process cannot continue. Two most common, and overlapping, varieties of telic processes are: human activities directed to a specific goal, and processes that consume a specific amount of a process-specific resource. So, reading a magazine article is a telic process not only because reading is a goal-directed activity, but also because an article contains a limited amount of the process-specific resource (reading matter) that gets steadily consumed as the process unfolds. Temporal information about a telic process thus includes a typical rate of progress towards its goal or a typical rate of consumption of its resource.

A comparison of the preparatory and the resulting stages of a telic process shows the possible changes that such processes bring about: a telic process can create an object, destroy an object, modify an object, or move a specified amount of material (possibly the mover himself) to a specified destination. A subclass of destruction processes are ingestions, which convert an external resource into an internal one. Moving is understood to include all three of Schank's PTRANS, ATRANS, and MTRANS classes (Schank 1975). Moving also includes gradual (but not instantaneous) changes of state.

It has been observed (in Verkyul 1972 and others after him) that the aspectual class of a telic predicate is

affected, in regular ways, by the properties of its noun-phrase (NP) complements, especially **theme**. Such distinctions as plural-singular, group-individual, and mass-count are particularly important, as well as the degree of granularity-homogeneity and internal structure. To take a well-known example, *Bill crossed the border* describes an instantaneous event; *The regiment crossed the border* describes a noninstantaneous h-token with a beginning, body, and end; and *Drug smugglers crossed the border easily* describes an activity that continues as long as there are drug smugglers. The aspectual properties of telic h-tokens thus arise compositionally in context;<sup>3</sup> however, it still makes sense to talk of the aspectual class of a telic h-type as it manifests itself in h-tokens with atomic individuals as themes.

#### 4.3 EVENTS, PROCESSES, AND STATES

Now I briefly comment on how the conceptual framework developed here relates to a classification of histories into events, processes, and states (Mourelatos 1981). For a linguist, the distinction **event-process** is one of aspectual perspective: "The term 'process' means a dynamic situation viewed imperfectly, and the term 'event' means a dynamic situation viewed perfectly" (Comrie 1976:51). The distinction **process-state** is one of aspectual class. Thus the sentence *John was walking to the store* describes a process, not a state, because walking in general, and walking to a specified destination in particular, is intrinsically a process; it describes a process and not an event because the h-token is viewed imperfectly. I think the lexicon can be made more consistent and informative if the notion of aspectual class (lexical aspect, Aktionsart) is kept separate from the aspectual sentence perspective.

An important piece of evidence for keeping the distinctions event-process and process-state separate is provided by events "made out of" states such as the event described by *Bobby took a nap*. (Cf. also *did some reading, went for a walk*.) Collocations of this nature have never, to my knowledge, been discussed in connection with the English aspectual system, precisely because, in English, it is easy to associate telicity (an intrinsic lexical property) and perfectivity (the value of a subjective grammatical category). In languages with a well-articulated perfective aspect a stretch of most atelic activity can be viewed perfectly, i.e., as an event. For instance, *He walked* (perfective), clearly atelic, would mean something like "He walked for a while/he took a walk"—clearly an event. Atelic histories, of course, lend themselves more readily to the imperfective view; Nakhimovsky 1983:80–2 documents some subtle ways in which "aspectual" languages distinguish between atelic events, which are events simply by virtue of allocating them a limited amount of time, and telic events, which are intrinsically bounded in time by the very manner in which they unfold. Telic events tend to be identified as a natural class because they

support perhaps the most common temporal inference we make: given a telic history viewed perfectly, we infer that the end point is reached. (This is made more precise in Section 4.5.) For all that, telicity and perfectivity remain distinct, if not entirely orthogonal, categories.<sup>4</sup>

#### 4.4 A CLASSIFICATION OF INSTANTANEOUS EVENTS

Unlike noninstantaneous histories, instantaneous events lack internal structure and have to be classified by comparing the world before and after them. An instantaneous event can terminate either a process or a state, and it can initiate either a process or a state; if it is sandwiched between two processes or two states, the two can be the same or different. The resulting classification is shown in Table 2, next page. It can be seen that in English, most verbs describing instantaneous events fall into the first three groups, where the instantaneous event meets a state. Sometimes the name of the process can be used to describe its beginning, as in *She ran*, i.e., "set out running."

Among the three groups of instantaneous events, happenings always have a progressive form, and they always describe a sequence of instantaneous h-tokens. Transitions may or may not have a progressive form; if they do, progressive is again associated with plurality, e.g., *He's been noticing that . . .* (the implication being that the noticing has taken place on many occasions. Observe the inference that between the occasions there is a return to the initial state of the transition: in order to again notice, i.e., become aware of something, one has to first cease to be aware of it.). Finally, culminations usually have progressive forms that allow two interpretations. One is the familiar idea of plurality: *The travelers, one by one, were reaching the summit*. The other arises from the nature of culminations, which crown processes with sharply defined end points: the top of a climb, the finish of a race. The progressive of a culmination verb describes an h-token whose lifetime is an open interval properly contained in the time of the process that leads to the instantaneous event. I conclude that the proposed classification, although based on extra-linguistic considerations, captures salient distinctions in the meaning of linguistic forms.

All the classifications developed in this section are brought together in tables 1–3, next page.

#### 4.5 ASPECTUAL CLASS, SENTENCE PERSPECTIVE, AND INFERENCE

Recall that the aspectual sentence perspective is determined by the position of the RT, either inside the time of an h-token or after that. In interaction with the knowledge about the internal composition of h-tokens, the position of the RT creates expectations, which can be stated using the notions of time scale and expended resources. Consider first an example: the h-token of Sharon reading a novel can terminate in one of the following four ways:

ASPECTUAL CLASSES

NOT POSSESSING INTERNAL STRUCTURE (instantaneous)	POSSESSING INTERNAL STRUCTURE	
	State	Process
	Atelic	Telic
SAME, VIEWED FROM INSIDE		

Table 1. Top-level classification.

Configuration	Label	Examples
state-E-same state	happening	flash, cough, jump
state-E-different state	transition	forget, recognize, notice,
process-E-state	culmination	win the race, reach the top
process-E-same process	disturbance	examples
state-E-process	activation	more difficult
process-E-different process	switch	to find

Table 2. Instantaneous events.

STATES		PROCESSES		
zero-resource	generic resource	atelic: generic resource	atelic: specific resource	telic
relations: own possess resemble	sleep stand sit	walk run think(of) work	read write build dig	read a book write a letter build a chair dig a hole shave
perceptions: see hear feel	lie hold lean against			
mental states: know remember trust				

Table 3. Noninstantaneous h-types.

**natural end:** Within a certain time scale, Sharon finished the novel;

**exhaustion:** Within a certain time scale, determined by a different set of resources, Sharon got tired of reading;

**volition:** Sharon decided to put the novel down;

**interruption:** Something happened that made reading impossible: Sharon's friend took the novel away from her, the lights went out, her attention was distracted by loud music.

Generalizing from this example, we can say that when the RT finds itself inside an h-token, some or all of the following expectations are set up, depending on the internal dynamics of the h-token and the presence or absence of volitional control:

- If the h-token has an Agent participant, it can be interrupted by the Agent's decision to do so. (The decision may turn out to be impossible to implement but the h-token is then transformed into a different one, in which the former Agent perhaps becomes an Object or Patient: think of a driver trying to stop the car and going into a skid.)
- If the h-token consumes generic resources (i.e., if it is a process or a resource consuming state) then, within

a certain time scale, it comes to an end when the generic resource is exhausted.

- If the h-token is a telic process, then, within a certain time scale, the process will come to its natural end.
- Finally, and most difficult to detect, the h-token may come to an end if some of its preconditions no longer hold, causing an interruption of the h-token.

The causality involved in interruption may be indirect, and a complete specification of everything that can go wrong is clearly impossible, but no such completeness is observed in human reasoning, either. What we can and ought to do is list in the lexicon the main preconditions for processes and states expressed by lexical items. Assuming that the lexicon is organized as an inheritance hierarchy, a great deal of preconditions will be shared by entire classes of histories. Some examples are given in Nakhimovsky 1987a.

The most noticeable expectations are those arising when a telic process is viewed imperfectly, and the logical problems of reasoning with expectations were first observed on such processes. This is Dowty's (1977) "imperfective paradox", i.e., the problem of giving "an account of how [John was drawing a circle] entails that John was engaged in bringing-a-circle-into-existence activity but does not entail that he brought a circle into existence" (p.46). From the perspective of the intervening years it can be seen that Dowty here grapples with the default, nonmonotonic nature of the inferences involved. The progressive sentence puts the RT in the middle of the time of John's activity; if another sentence puts the RT after that interval, we infer, unless or until told otherwise, that the activity came to its natural end. More precisely, the inference works as follows:

**Telic-Imperfective inference rule:** If one sentence in a narrative tells us that a telic history h1 has begun or is in progress, and a later sentence moves the RT into or beyond a later history h2, and the time distance between h1 and h2 is of the same or a greater scale than the time scale of h1, then assume, unless or until told otherwise, that the end-point of h1 is reached.

The rule can be illustrated by the following narrative: *In the afternoon, when Johnny was writing a postcard to his friend, Mommy went to the market. In the evening they played together.* Knowing how long it takes, we infer that Johnny finished his postcard and Mommy her shopping. However, the inferences are withdrawn if contradicted: the narrative could continue with *until Mommy discovered that the postcard never got written.* To see the importance of time scale, replace "writing a postcard" with "reading *War and Peace.*"

4.6 DURATIONAL CONSTRAINTS

The examples of the preceding section illustrate two facts about temporal reasoning that I believe are crucial for any implementation of narrative understanding: (a) knowledge about internal constituency of events and temporal relations between them is tightly integrated



with durational knowledge; (b) durational knowledge includes a great many constraints originating in celestial, biological, and civilization-specific cycles. The first of these facts is recognized and implemented in Allen and Kautz 1985. The implementation suggestions made here all derive from that paper and Forbus 1985. Allen and Kautz also note that “in more complicated situations, certain units (such as the standard time units—minutes, hours, weeks, etc.) can be used as *reference durations* to keep the network of manageable size” (p. 260). How this ties in with (b) above can be seen by recasting Allen and Kautz’s example of a unified (temporal relations + durations) implementation in the present framework. The example in its entirety is as follows (pp. 260–1):

**Example 5**

- a. Moe and Larry began reading Principia Mathematica.
- b. Moe read for over an hour.
- c. Larry stopped reading and fell asleep after 10 minutes.

Allen and Kautz interpret (5a) to mean that Moe and Larry started reading simultaneously (and may or may not have finished together). But surely, language is not that precise: (5a) can be said about two students who both started reading the book in September for a course they were taking in the fall semester. The most one can say about (5a) is that the time distance between [began-read Moe] and [began-read Larry] is **much-shorter-than** the expected duration of reading. This is an instance of a general rule, noted in the end of Section 2, that the duration of an h-token is **much-greater-than** the durations of its beginning and end.

Sentence (5b) supplies both a duration and a time scale. Allen and Kautz’s representation again under-constrains: they interpret the sentence to mean that the duration of Moe’s reading is  $(1 \infty)$ . An obvious constraint is that Moe read for less than a day, but more precisely “over an hour” implicates “less than two hours”; in general, “over/after n units of time” implicates “less than n + 1 units.” (This is an instance of Horn’s (1984) scalar implicature.) Allen and Kautz, in fact, make an implicit use of this implicature when they interpret (5c) to mean that Larry read 10 to 15 minutes (p. 261). The interpretation is correct because 5 minutes is a duration-measurement unit in our language, as mentioned in Section 3 above.

It appears that the varieties of temporal knowledge proposed in this paper result in a richer set of constraints that give a more adequate picture of (5) and can handle more complex narratives. They also indicate a fairly clear connection between the task of natural language understanding and the enterprise of naive or qualitative physics (Hayes 1985, 1985a; Forbus 1985). The entire aspectual classification of histories is based on taking seriously the physical vocabulary of processes and states, and the reasoning of the preceding section is

an example of Forbus’s (1985) limit analysis. In general, nonmonotonic constraint propagation over interacting processes evolving in time is the essence of naive physics. These similarities suggest that qualitative process descriptions of Forbus 1985 can serve as a model for lexical entries describing h-types. (Some examples can be found in Nakhimovsky 1987a.)

## 5 NARRATIVES AND THEIR REPRESENTATIONS

Discussions of tense and aspect in discourse invariably seem to converge on **narrative discourse**, rather than conversation or exposition. The reasons, I believe, have to do with several profound, if somewhat speculative, distinctions: between semantic and episodic memory; between on-line **perception** of time effected by biological clocks and detached time **cognition** rooted in long-term memory; and between subjective and objective discourse. Briefly and, I repeat, speculatively, exposition contributes to semantic memory while narrative builds an episodic representation; conversation is a breed apart because the time of a conversation coincides with the time of its content, with simultaneity in time typically accompanied by a tight integration between the linguistic and nonlinguistic behavior: the verbalization of how-to-get-there directions *is* the action of giving directions; task-oriented conversations between an expert and apprentice are an integral part of performing the task at hand, and the unfolding *text* of an argumentative dialogue is precisely the activity of arguing. Conversation can thus properly be called **performative discourse**. By contrast, the content of a narrative is decoupled from the linear progression of its text and unfolds in its own, separate timeline.<sup>5</sup>

Whatever the reasons, it is certainly the case that (a) tense, aspect, and mood systems of the verb are frequently more finely nuanced in the past than in the present or future, and never the other way around; (b) the tense of narrative discourse is past, in its many varieties; (c) the past tense in narratives is not interpreted with respect to the situation of discourse. In place of the situation of discourse, a narrative is processed with respect to a constantly maintained **deictic center**, which is “the locus in conceptual space-time of the objects and events depicted or described by the sentences currently being perceived. At any point in the narrative, the cognitive agent’s attention is focused on particular characters (and other objects) standing in particular temporal and spatial relations to each other. Moreover, the agent ‘looks’ at the narrative from the perspective of a particular character, spatial location and temporal location. Thus the deictic center consists of a WHERE-point, a WHEN-point and a WHO point.” (Bruder et al., 1986:1). In this paper, the WHEN-point of the deictic center is referred to as the **temporal focus**, or TF (cf. Webber 1987a,b; Nakhimovsky 1987b). This term is adopted because it brings out the similarity

between definite and temporal anaphora (cf. Webber, this volume, for the genesis of this idea.).

Not one but two structures are thus needed to represent a narrative: the **event-situation structure** (ESS, cf. Webber 1987b), representing the narrative's unfolding contents, and the **linear text structure** (LTS), whose components are linked by rhetorical relations such as elaboration, resumption, or flashback (see, e.g., Hobbs 1982). Next to ESS and LTS, there must also be a short-term memory structure called **current focus space** (CFS) and a buffer in which the representation of the current sentence is assembled. Among the components of the CFS are the data structures needed for comprehension of definite anaphora (such as the Focus Stack and Alternate Focus List in Sidner 1983), as well as the components of the current deictic center, which are really pointers to the appropriate nodes of the ESS. After the first sentence of a narrative establishes the initial set-up, each subsequent sentence is processed in the context consisting of the ESS and the CFS. Depending on the sentence's meaning, one of two things happen: either the representation of the sentence is incorporated in the CFS, with the focusing mechanisms appropriately modified; or, in the case of a focus shift, the contents of the CFS are incorporated into the ESS and LTS, and the CFS is completely reset. In terms of text structure, the current sentence either continues the same, or starts a new, **discourse segment** (DS; cf. Grosz and Sidner 1986).

The nature of the processing at the DS juncture is thus quite different from the "routine" tasks to be performed as long as the text remains in the same DS: the start of new DS prompts, and is prompted by, a shift of attention. The circularity here is deliberate. The start of a new DS brings about several changes, some of them more immediately noticeable than others; the more obvious ones serve to signal that a new DS is, indeed, started. Some of the most important such signals come from changes in time scale, aspectual class, and other temporal characteristics of narrated h-tokens. These temporal discontinuities are usually accompanied by discontinuities of other kinds that form recurrent patterns; I suspect, and this is an empirical hypothesis, that rhetorical relations within a narrative text are macrolabels that stand for certain oft-repeated clusters of discontinuities. In any event, the task of segmenting narratives into units that form the basic elements of discourse representations and constrain the application of focusing algorithms is obviously very important. The task is frequently mentioned in work on discourse (see Grosz and Sidner 1986:177–8 and references there), but the available heuristics mostly rely on clue words and phrases. I propose heuristics based on four kinds of discontinuities: discontinuities of topic, discontinuities of space and time, discontinuities of figure and ground, and discontinuities of the narrative perspective (not to be confused with aspectual sentence perspective). I

explain these after a preliminary and partial illustration. Consider Example 6.

#### Example 6

- a. Hartley and Phoebe had been sent by their mother to fix the tail vane of the windmill.
- b. In the great expanse of the prairie where they lived, the high tower of the windmill was the only real landmark. (Worline 1956:1)

Rhetorically speaking, the (b) sentence above interrupts a sequence of events (DS1) to start DS2, a description. In order to recognize this rhetorical relation between the two DSs it is necessary to recognize that:

- a. there is a shift of topic;
- b. there is a shift in perceptual modality to visual perception;
- c. there is a shift in time scale from the events of the current day to years or decades, associated with the lifetime of a windmill and the 'where they lived' clause;
- d. there is a shift in spatial scale from a household to the entire prairie;
- e. there is a shift from a foregrounded sequence of events to the 'mopping-up' operation of filling the background. The shift is signaled by changes in aspectual class.

The following empirical investigation thus suggests itself: classify the discontinuities and clusters of discontinuities that typically accompany DS breaks in narratives; identify the linguistic and extralinguistic knowledge involved; develop heuristics for using this knowledge; and test the heuristics in a computer program.<sup>6</sup> This paper presents a classification that arises from an analysis of what constitutes a narrative.

The first observation to make is that a narrative must have a **plot**, i.e., present a sequence of events that forms an instance of a recognizable pattern. (The patterns are part universal, part culture specific; the work in Lehnert 1982 can be seen as a search for the principles on which such patterns are built.) Using the Gestalt terminology (brought into linguistics by Talmy 1983) I can say that a narrative's plot must present a recognizable temporal/causal figure shown against some ground that minimally consists of spatial/visual settings (descriptions of characters are also frequent). The distinction is not always clear-cut because elements of the figure can be hidden among the details of the ground, but the temporal nature of the plot does stand in clear contrast to the spatial nature of the ground.

Secondly, a narrative must have characters with which we empathize. These characters don't have to be human: one can easily imagine a story about an adventuresome plant seed that falls off its parent, gets swallowed and excreted by a horse, and nearly drowns in a tropical rain before being miraculously saved by the sun and producing a flower. Even so, the narrative is likely to alternate between the objective narrator's point of

view and that of one of the characters. ("The belly of the horse was dark and noisy inside.") Some discontinuities of narratives very simply reflect the intrinsic, spatial, or temporal discontinuities of the plot, when, e.g., the story is composed of a sequence of events taking place during an afternoon, followed by two years' hiatus, followed by another action-packed afternoon. (A similar example of a spatial discontinuity can be easily imagined.) One measure of the "simplicity" of a narrative is how faithfully the order and structuring of its text reflects the order and structuring of its component events. Even in the simplest narrative, however, there are bound to be discontinuities resulting from the tension between the linear nature of the text and the multidimensional structure that it is meant to evoke. These are discontinuities of figure and ground, when the narrative shifts between the main story line and the surrounding circumstance, and discontinuities of perspective, when the narrative crosses into a different "empathy space" or creates a new one.

Given this classification of discontinuities one can proceed to catalogue the clues that signal them. This is a subject for a large empirical study, of which the next section is but a preliminary sketch. It is important to keep in mind that it is *clusters* of discontinuities that signal the beginning of a new DS.

## 6 DISCONTINUITIES IN NARRATIVES

### 6.1 DISCONTINUITIES OF TOPIC

Discontinuities of topic fall in two groups. In the first, there is no anaphoric relation or immediate inference path from the new topic to a node in the CFS. What "an immediate inference path" is depends, of course, on the system's knowledge base and inferential capabilities, but this is a separate big issue that is not dealt with in this paper. The second kind of discontinuity, noted in Reichman 1985, arises when an anaphoric relation exists and allows a pronominal anaphor, but the topic is instead reintroduced by a full noun phrase. I call this phenomenon **topic reintroduction** and illustrate it in 6.5.

### 6.2 TEMPORAL DISCONTINUITIES

The most important temporal discontinuities are:

- (a) A shift from perfective to imperfective sentence perspective accompanied by a shift to a much greater time scale. The corresponding rhetorical move is frequently characterized as "introducing background or descriptive material". The move is frequently accompanied by a topic reintroduction.
- (b) The reverse shift from descriptive material to the main line of the narrative. This move is signaled by the **temporal focus** (TF), and the entire deictic center, returning to an established node in the ESS, with an appropriate contraction of the time scale.
- (c) A backwards move of the TF to an earlier point in

time, with or without a change in time scale. Rhetorically, this is known as flashback. This move is frequently signaled by a verb in past perfect or by the *used to* + infinitive construction, although a shift may occur without such a verb form, and the presence of such a verb form does not necessarily signal a shift: the reference time of the sentence may remain the same as, rather than precede, the current TF. (Section 7 illustrates; see also Almeida 1987 and Webber 1987b, this volume.)

### 6.3 SPATIAL DISCONTINUITIES

The only purely spatial discontinuities I have looked at are discontinuities of scale. Just as h-types have time scales associated with them, characterized in terms of "received" cyclical events such as day or year, it seems equally necessary to establish a gradation of spatial scales, based on similar considerations from human biology and habitat. (A hint at the same idea can be found in Nirenburg et al., 1985:230.) Some important spatial scales are: tiny, hand, one's body (may need to be further specialized); within arm's reach; room area, such as desk or bed; room; floor; house; household; village or neighborhood; larger area within a between-meals round trip; within a day's round trip; staying overnight. (The larger the scale, the more domain- and culture-specific variation there is.)

### 6.4 DISCONTINUITIES OF PERSPECTIVE

Two kinds of phenomena are clearly relevant: attitude reports, and indications of the spatial position from which things are perceived (the WHO-point in Bruder et al., 1986). The subject needs further study.

### 6.5 AN EXAMPLE

The remainder of this section shows how the above heuristics work themselves out on the initial fragment of Joyce's "Clay" (Joyce 1969:99–106). The fragment falls into three DSs: Sentence 1, sentences 2–6, and sentences 6–12; Sentence 13 starts a fourth DS by returning to the first one. Embedded in DS3 is DS3.1 consisting of sentences 9–10.

1. The matron had given her leave to go out as soon as the women's tea was over and Maria looked forward to her evening out.
2. The kitchen was spick and span: the cook said you could see yourself in the big copper boilers.
3. The fire was nice and bright and on one of the side-tables were four very big barmbracks.
4. These barmbracks seemed uncut; but if you went closer you would see that they had been cut into long thick even slices and were ready to be handed round at tea.
5. Maria had cut them herself.
6. Maria was a very, very small person indeed but she had a very long nose and a very long chin.
7. She talked a little through her nose, always soothingly: *Yes, my dear*, and *No, my dear*.
8. She was always sent for when the women quarrelled over their tubs

and always succeeded in making peace. 9. One day the matron had said to her:—Maria, you are a veritable peace-maker! 10. And the sub-matron and two of the Board ladies had heard the compliment. 11. And Ginger Mooney was always saying what she wouldn't do to the dummy who had charge of the irons if it wasn't for Maria. 12. Everyone was so fond of Maria.

13. The women would have their tea at six o'clock and she would be able to get away before seven.

The transition from Sentence 1 to 2 is very similar to the transition from (a) to (b) in Example 6: it is a shift from figure to ground marked primarily by aspectual changes and a shift from temporal to spatial/visual material. The change in time and space scales is not as dramatic here, but, on the other hand, there is a discontinuity of perspective as the story shifts from Maria's attitude report ("Maria looked forward") to the objective perspective of the implied narrator ("The kitchen was spick and span . . .").

The transition from 5 to 6 (paragraph break) is characterized by a shift in time scale and a topic reintroduction. The material is a (back)ground character description: "Maria is such that. . . ." The beginning of DS3.1 is signaled by a well-known "clue phrase" *One day* and by the past perfect tense. Note that the material is still ground ("Maria is such that. . . .") but the precise relationship between DS3 and DS3.1 is ambiguous: they can be siblings. Only when DS3 returns in Sentence 11 (the signals are the return to the simple past and the sentence-initial *and*) is it established that DS3.1 is, indeed, embedded in DS3.

The transition from 12 to 13 (the next paragraph break) is a return to the event sequence of Sentence 1. The position of the TF, and the entire deictic center is indicated by the future-in-the-past tense. Note that it is essential to retrieve the entire deictic center, and not just the TF, because the WHO-point is also restored: the *she* in 13 does not evoke *Maria* of Sentence 12 or any other sentence in DS2 and 3 presented from the perspective of the implied narrator. Rather, this pronoun is a quasi-indexical (Rapaport 1986) that replaces the first-person singular *I* of a portion of the narrative from Maria's perspective, signaled by *Maria looked forward* in Sentence 1.

## 7 NARRATIVE MOVES

It is a common intuition that, as a narrative progresses, TF changes its position in time. I will refer to the movement of TF from one sentence of the narrative to the next as the **narrative move**. Narrative moves can be classified along several dimensions. First off, the TF can remain in the same timeline or branch off to an alternative one. (Such branchings-off are usually signaled by a nonindicative mood of the verb; this paper has nothing more to say about them.) Within the same timeline, the TF can move forward, sideways (into an

overlapping h-token), or backwards, to an earlier h-token. In each of these three cases, the move may or may not be accompanied by a change in time scale. Finally, the TF can remain within the current focus space or shift to a different one. (In terms of linguistic structure, the current sentence may continue the same, or start a new, DS.) I will coin the terms **micromove** and **macromove** for the last distinction.

The preceding section argued that, if considered in the context of the entire deictic center movements, macromoves can be fully catalogued and associated with a limited set of rhetorical labels. It appears even more likely that micromoves also form a small set of recurrent patterns, because they are additionally constrained by the requirement that they cannot introduce a drastic discontinuity. I present a sample of micromoves below, contrasting them with macromoves; the exposition is organized by the temporal classification of narrative moves (backward, sideways, or forward). Much of the material of this section overlaps with Webber (this volume); some of the examples are from Dowty 1986.

### 7.1 BACKWARD MOVES

Backward moves are always macromoves, i.e., they always introduce a DS break. This is not as drastic a claim as it might seem, because simply narrating an h-token that precedes the current TF does not necessarily introduce a backward move. In particular, one sentence in the past perfect does not produce a backward move. (This point is made, with respect to the French plus-que-parfait, in Kamp and Rohrer 1983: 256 fn.) If the out-of-sequence h-token is presented perfectly, i.e., the RT is after it, the TF does not move backward. Compare the following two examples:

#### Example 7

The telephone rang: it was Mme Dupont; her husband had eaten too many oysters for lunch. The doctor recommended a change in lifestyle.

#### Example 8

. . . Tessie sat beside [Phoebe] with Araminta. Araminta had begun as an ear of corn. On a piece of cloth Mother had painted a face, using bluing for the eyes . . . She had stitched on smaller pieces of cloth, black for the nose, pink for the checks and red for the mouth. She had tied the cloth over one end of the ear of corn, then made a bonnet to go over that . . . All in all Araminta was one of the most elegant members of the Dawson family. (Worline 1956:28).

In (7) the past perfect is like the present perfect in a dialogue: it indicates precedence and relevance but does not shift the TF. In (8) the past perfect is like a past tense in a dialogue: it creates a new TF position to which subsequent sentences refer, until simple past

terminates the embedded DS and returns to the one interrupted by the backwards move.

A sentence in the simple past following a sentence in the past perfect does not always terminate an embedded DS: nontemporal anaphora may integrate the two sentences (cf. Webber, Example 20 in this issue, and discussion). Such a past perfect/simple past configuration is a must if one wants to go into a long, multi-FS digression, because there seems to be a strong limit on how long an embedded DS can continue in the past perfect, as in (8). Consider:

**Example 9**

John was the captain of Penn's squash team. He had previously been captain of the Haverford team, but the coach there started chasing after John's girlfriend and John got very upset. One day he went to talk to the coach, but . . .

Example 9 can now continue with an embedded narrative of arbitrary length. However, the interrupted DS has to be eventually resumed, otherwise the narrative is perceived as sloppy (or, if Proust does it, self-conscious, cf. Genette 1980:66–67).

**7.2 SIDEWAYS AND FORWARD MOVES**

The majority of micromoves belong in these two categories, and in this brief subsection I cannot possibly do justice to them all. What I try to do is establish a framework, based on the conceptual apparatus of this paper, for a more detailed investigation.

I assume, for simplicity, that both the most recently processed sentence (S0) and the current sentence (S1) consist of just one clause each, narrating h-tokens H0 and H1, respectively; the verbs in the sentences, V0 and V1, can be either past or progressive, and their aspectual classes can be any of those introduced in Section 3. Consider a few combinations.

**H0: telic, simple past-H1: telic, simple past**

**Example 10**

- John entered the president's office. The president got up.
- John blew his nose. The president mopped his brow.
- John hid in the president's office. Bob locked himself up in the closet.

This combination is most frequently a sequence, as in (a), but not necessarily so. If H0 and H1 are not related by an obvious chain of causality the sequencing is not obvious, and if they both can be construed as responses to the same event, they are interpreted as overlapping, cf. (b) and (c).

**H0: telic, simple past-H1: atelic or state, simple past**  
**H0: telic, simple past-H1: non-instantaneous, progressive**

**Example 11**

- John entered the president's office. The president was asleep.
- John entered the president's office. The president was writing a letter.

In these examples, there is no compositional relation between H0 and H1, and H1 is interpreted as beginning before H1 and continuing after it. However, consider the following texts, in which H1 is part of the body, or the resulting stage, of H0:

**Example 12**

John knelt at the edge of the stream and washed his hands and face. He washed slowly, feeling the welcome sensation of the icy water on his parched skin.

**Example 13**

- Bob turned off the light. It was dark around him.
- Bob heard a loud click. The door behind him was locked.
- Bob closed his eyes. He was asleep.

In (13), the narrative moves forward, into a narrated h-token. The conditions for the move are that there is a direct inferential link from H0 to H1, and/or some sort of incompatibility between them. So, before Bob's eyes were closed they had to be open, which, for most of us, is incompatible with being asleep; on the other hand, one frequently closes one's eyes with the intention of going to sleep.<sup>7</sup>

**H0: state, atelic or progressive-H1: state, atelic or progressive**

**Example 14**

- John was fixing an electrical appliance downstairs. Suddenly, the whole house was dark.
- The president was conducting the interview with Mary in a cold, official manner. Suddenly, to her utter amazement, he was offering her an important job.

The contrast between (11) and (13) illustrates the following narrative principle: it is easy to move *sideways* into a narrated h-token, but it is difficult to move *forward* into one. (In other words, an imperfective sentence will not be interpreted as advancing the TF unless such an interpretation is forced by some form of incompatibility and causality.) It is especially difficult to move forward *from inside* one interval into another one, as in Example 14. In fact, such a move cannot be effected without the help of an adverbial that acknowledges the gap (try removing *suddenly* from the examples).

This concludes my discussion of narrative moves. The sampling given here is certainly not complete, but an extensive empirical investigation can, I believe, produce a full catalogue, organized by the configuration of tenses, aspectual classes, and time scales of the

h-tokens involved. As the examples of this section show, no such configuration determines the narrative move uniquely, but they strongly constrain the choices.

## 8 CONCLUSIONS

This paper suggests several empirical investigations having to do with the lexicon and discourse. In conclusion, I would like to recapitulate them and outline some possible ways to proceed.

In order to enter durational information into the lexicon, the time-measurement units need to be identified, from **moment**, **minute**, and **five minutes to year** and **decade**. Then, for many, perhaps most, h-types, their time scales can be characterized by a sequence of such units, e.g., *visit-1* [**minutes, hours**] “visit somebody” (as opposed to *visit-2* [days] “visit with somebody”). Two h-types stand in the **much-greater-than** relation to each other if their time scales do not overlap.

Aspectually, each h-type should be assigned to a class according to the classification developed in Section 4. Regular aspectual polysemies should be noted and their extent investigated. It would be helpful to know, for example, how many, and what kind of, process verbs can describe the instantaneous beginning of the process as in *She ran* (i.e., began to run). For telic h-types, as argued in Section 4, there is a limited number of ways in which the arguments of the h-type can be affected by it (**create, destroy, move, modify, expend**: see Section 4.2). These ways can be catalogued and used as labels on the appropriate slots of lexical representations. Qualitative dependencies between arguments can be described using Forbus’s (1985) approach. (Nakhimovsky 1987a has examples.) A lexicon so outfitted can be used in a constraint propagation program, similar to the Allen and Kautz 1985 experiment, in a manner described in Section 4.6.

With respect to narratives, two investigations are in order, one concerning micromoves within a DS, the other concerning macromoves and DS boundaries. The aim of the former investigation would be to identify the most frequent temporal/causal patterns of micromoves and how they are signaled by tense and aspect. The aims of the latter investigation would be, first, to identify the types of discontinuities that cause the shift to a new DS; second, identify typical combinations of such discontinuities and how they correspond to rhetorical notions and the narrator’s intentions; and finally, develop heuristics for the proper segmentation of narratives. At this stage, work outlined here is likely to connect with some of the pursuits of structuralist and poststructuralist poetics.

## REFERENCES

- Allen, James 1984 Towards a General Theory of Action and Time. *Artificial Intelligence* 23(2): 123–54.
- Allen, James and Kautz, H.A. 1985 A Model of Naive Temporal Reasoning. In Hobbs and Moore 1985.
- Apresjan, Juri 1974 *Lexical Semantics*. (In Russian) Nauka, Moscow.
- Almeida, Michael J. 1987 Reasoning About the Temporal Structure of Narratives. Ph.D. diss. and TR 87–10. Computer Science Department, SUNY at Buffalo, Buffalo, NY.
- Bach, Emmon 1986 The Algebra of Events. *Linguistics and Philosophy* 9: 5–16.
- Bobrow, Daniel, (ed.) 1985 *Qualitative Reasoning about Physical Systems*. MIT Press, Cambridge, MA.
- Bruce, Bertram 1981 A social interaction model of reading. *Discourse Processes* 4: 273–311.
- Bruder, G.A. et al. 1986 Deictic Centers in Narrative: An Interdisciplinary Cognitive-science Project. TR 86–20. Computer Science Department, SUNY at Buffalo, Buffalo, NY.
- Bunt, H. 1985 The Formal Representation of (quasi) Continuous Concepts. In Hobbs and Moore 1985.
- Campbell, Jeremy 1986 *Winston Churchill's Afternoon Nap*. Simon and Schuster, New York, NY.
- Comrie, B. 1976 *Aspect*. Cambridge University Press, Cambridge, England.
- Cooper, Robin 1986 Tense and Discourse Location in Situation Semantics. In Dowty 1986a: 17–36.
- Creary, L.; and Pollard, Carl 1985 A Computational Semantics for Natural Language. *Proceedings of the 23rd Annual Conference of the Association for Computational Linguistics*, Chicago, IL; 172–9.
- Dahl, O. 1985 *Tense and Aspect Systems*. Basil Blackwell, Oxford, England.
- Dowty, David 1986 The Effects of Aspectual Class on the Temporal Structure of Discourse. In Dowty 1986a: 37–61.
- Dowty, David, (ed.) 1986a Tense and Aspect in Discourse. *Linguistics and Philosophy* 9(1) (special issue).
- Dowty, David 1987 Aspect and NP semantics. *Logic and Linguistics Conference*. Stanford, CA.
- Dowty, David 1977 Toward a Semantic Analysis of Verb Aspect and the English “imperfective” Progressive. *Linguistics and Philosophy* 1: 45–77.
- Forbus, K.D. 1985 Qualitative Process Theory. In Bobrow 1985.
- Fraisse, Paul 1984 Psychology of Time Perception. *Annual Review of Psychology* 1–34.
- Genette, G. 1980 *Narrative Discourse*. Cornell University Press, Ithaca, NY.
- Grosz, Barbara; and Sidner, Candace 1986 Attention, Intentions, and the Structure of Discourse. *Computational Linguistics* 12(3): 175–204.
- Hayes, P. 1985 The Second Naive Physics Manifesto. In Hobbs and Moore 1985.
- Hayes, P. 1985a Naive Physics I; Ontology for Liquids. In Hobbs and Moore 1985.
- Heim, Irene 1982 *The Semantics of Definite and Indefinite Noun Phrases*. PhD diss. University of Massachusetts, Amherst, MA.
- Hinrichs, Erhard 1986 Temporal Anaphora in Discourses of English. In Dowty 1986a: 63–82.
- Hirst, G. 1987 *Semantic Interpretation and the Resolution of Ambiguity*. Cambridge University Press, Cambridge, England.
- Hobbs, J.R. 1982 Towards an Understanding of Coherence in Discourse. In Lehnert and Ringle (eds.). *Natural Language Processing*. Erlbaum Associates, Hillsdale, NJ.
- Hobbs, J.R. 1985a Granularity IJCAI 85: 432–5.
- Hobbs, J.R. and R. C. Moore (eds.) 1985 *Formal Theories of the Commonsense World*. Ablex Publishing Corporation, Norwood, NJ
- Hobbs, J.R. et al. 1985 *Commonsense Summer: Final Report* (no. CSLI-85-35) CSLI, Stanford, CA.
- Hobbs, J.R. et al. 1986 Commonsense Metaphysics and Lexical Semantics. *Proceedings of the 25th Annual Conference of the Association for Computational Linguistics*, Stanford, CA; 231–40.

- Horn, Larry 1984 Toward a New Taxonomy for Pragmatic Inference. In D. Schiffrin (ed.) *Meaning, Form and Use in Context: Linguistic Application*. Georgetown University Press, Washington, DC.
- Joyce, James 1969 [1914] "Clay." In *Dubliners*. Modern Library, New York, NY; 99–106.
- Kamp, Hans 1981 A Theory of Truth and Semantic Representation. In J.A.G. Groenendijk, et al. (eds.), *Formal Methods in the Study of Language*. Amsterdam.
- Kamp, Hans and Christian Roerer 1983 Tense in Texts. In R. Bauerle, C. Schwarze, Arnim von Stechow (eds.), *Meaning, Use and Interpretation in Language*. Walter de Gruyter, Berlin, Germany.
- Lehnert, Wendy 1982 Plot Units: A Narrative Summarization Strategy. In W. Lehnert and M. Ringle (eds.), *Strategies for Natural Language Processing*. Erlbaum, Hillsdale, NJ.
- Link, G. 1983 The Logical Analysis of Plurals and Mass Terms: A Lattice-theoretical Approach. In R. Bauerle et al. (ed.), *Meaning, Use and the Interpretation of Language*. Walter de Gruyter, Berlin and New York.
- McDermott, Drew 1982 A Temporal Logic for Reasoning about Processes and Plans. *Cognitive Science* 6: 101–155.
- Moens, M.; and Steedman, M. 1987 Temporal Ontology in Natural Language. In *Proceedings of the 26th Annual Conference of the Association for Computational Linguistics*, Buffalo, NY; 1–7.
- Mourelatos, A.P.D. 1981 Events, Processes, and States. In Tedeschi and Zaenen.
- Nakhimovsky, A. 1983 *Meaning-Text Linguistics and the Problem of Voice*. Linguistic Research Inc., Edmonton, Alberta.
- Nakhimovsky, A. 1986 The Lexicon and Commonsense Reasoning: A Case Study. *Proceedings of the Conference on Advances in Lexicology*. University of Waterloo, Waterloo, Ontario.
- Nakhimovsky, A. 1987a The Lexicon, Grammatical Categories, and Temporal Reasoning. In J. Hallam and C. Mellish (eds.), *Advances in Artificial Intelligence*, John Wiley, Chichester, England.
- Nakhimovsky, A. 1987b Temporal Reasoning in Natural-Language Understanding: The Temporal Structure of the Narrative. EACL-87. Copenhagen.
- Nakhimovsky, A. and Rapaport, W. In preparation. Discontinuities in narratives.
- Nirenburg, Sergei; Raskin, Victor; and Tucker, Allen 1985 Interlingua Design for TRANSLATOR. *Proceedings of the Colgate Conference on Machine Translation*. Hamilton, NY.
- Passonneau, Rebecca. 1987 Situations and Intervals. In *Proceedings of the 26th Annual Conference of the Association for Computational Linguistics*, Buffalo, NY; 8–15.
- Rapaport, William J. 1986 Logical Foundations for Belief Representation. *Cognitive Science* 4: 371–422.
- Reichman, R. 1985 *Getting Computers to Talk Like You and Me*. MIT Press, Cambridge, MA.
- Schank, R. (ed.) 1975 *Conceptual Information Processing*. North-Holland, Amsterdam.
- Sidner, Candace 1983 Focusing in the Comprehension of Definite Anaphora. In M. Brady and R. Berwick (eds.), *Computational Models of Discourse*. MIT Press, Cambridge, MA.
- Talmy, Leonard 1983 *How Language Structures Space*. Berkeley Cognitive Science, Report no. 4, Berkeley, CA.
- Tedeschi, P.; and Zaenen, Anne (eds.) 1981 *Tense and Aspect* (Syntax and Semantics 14), Academic Press, New York, NY.
- ter Meulen, Alice 1987 Structured Domains for Events. *Workshop on Generics and Mass Terms*. Stanford, CA.
- Tompkins, Jane 1981 *Reader Response Criticism*. Johns Hopkins University Press, Baltimore, MD.
- Vendler, Z. 1967 Verbs and Times. In *Linguistics and Philosophy*. Cornell University Press, Ithaca, NY.
- Verkuyl, Henk 1972 *On the Compositional Nature of Aspect*. Reidel, Dordrecht.
- Webber, Bonnie 1979 *A Formal Approach to Discourse Anaphora*. Garland, New York, NY.
- Webber, Bonnie 1987a Two Steps Closer to Event Reference. TR MS-CIS-86-74, Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA.
- Webber, Bonnie 1987b The Interpretation of Tense in Discourse. In *Proceedings of the 26th Annual Conference of the Association for Computational Linguistics*, Buffalo, NY; 147–154.
- Wierzbicka, Anna 1985 *Lexicography and Conceptual Analysis*. Karger, Ann Arbor, MI.
- Worline, Bonnie Bess 1956 *The Children Who Stayed Alone*. Scholastic Book Services, New York, NY.

## NOTES

1. I am grateful to Rebecca Passonneau for detailed comments on an earlier version, and to Bonnie Webber for many discussions and painstaking editorial work. Conversations with Wayles Browne, Donka Farkas, Tom Myers and Mark Steedman have been helpful in clarifying my views (which is not to claim that clarity has been achieved). Patricia Ryan of Colgate has made it possible for me and the deadline to meet. Work on the paper was mostly done during a study leave granted by Colgate.
2. It was pointed out to me (cf. also Comrie 1976) that modern technology, and in particular slow motion movies, can spread out instantaneous events so that one can imagine somebody saying *Look, he's blinking!* or *He's beginning to blink!* in reference to a protracted individual blink on the screen. I don't think this abolishes the durational basis for the notion of instantaneous events: they're still treated differently than noninstantaneous ones, and the above sentences refer to an event on the screen, not a real-life blink. We wouldn't want to say that just because we can produce a huge image of a mosquito we no longer think of it as something small, or that the phrase *No larger than a mosquito* can no longer be used.
3. See Bach 1986, ter Meulen 1987. Dowty 1987 reports an idea of Hinrichs' that "with respect to their Theme arguments, the meanings of telic predicates are homomorphisms from the algebra of entities (NP denotations) [as suggested in, e.g., Link 1983—AN] into the algebra of events (as conceived by Bach)."
4. Dahl (1981) confuses the issue by defining perfectivity only for telic histories (p. 82), in terms of reaching or not reaching their terminal point. Not surprisingly, he finds it difficult to keep the two categories apart.
5. Clearly, both narrative and exposition can be embedded in a conversation: the modality (spoken or written) does not determine the genre uniquely. I am talking about clear-cut cases.
6. One such investigation is already under way as part of a larger interdisciplinary study of the Deictic Center in Narratives (Bruder et al., 1986). Some of the foregoing material has been influenced by that project and overlaps with Nakhimovsky and Rapaport (in preparation).
7. Discussing examples like these, Cooper 1986:34 says that this kind of move is impossible with progressives. I don't think this is the case; once the dynamics of the move is understood, an appropriate example is easy to construct, e.g.: A heavy volume slipped off the top shelf and landed on Bob's right toe. He was hopping on the unaffected foot, holding the hurt one in both hands.