PARALLEL EVENT STRUCTURES: EVIDENCE FROM ENGLISH

KOICHI MIYAKOSHI
University of Tsukuba

The present study is concerned with the proper representation of events, and argues for a new model of event representation, which is referred to as the “parallel event structure (PES).” The PES plots time on the horizontal dimension, and has two or more tiers aligned on the vertical dimension, each of which encodes a certain type of event (e.g. action/change). This multi-tiered structure represents three types of information on the two dimensions: it mainly represents (i) force-dynamic information (the exertion of energy and its effect) vertically, (ii) simple aspectual information (the temporal evolution of each subevent) horizontally, and (iii) complex aspectual information (two or more subevents’ parallel unfolding over time) two-dimensionally. It also represents the flow of time and the unfolding of (the subevents of) an event in terms of scale, thereby capturing their homomorphism as well as parallelism. This article argues that the PES is preferable to alternative models of event representation in terms of descriptive power and uniformity/coherence, showing that it clearly delineates various types of events from one another while capturing their commonalities in a uniform and coherent way. Empirical evidence is adduced primarily from verbal and adverbial expressions in English. Theoretical issues revolving around event structure are also discussed, which include force-dynamics, aspect, telicity, and scale structure.*

Keywords: parallel event structure, action/change tier, force-dynamics, aspect, telicity, scale

1. Introduction

The vast amount of work on lexical semantics and syntax over the past forty years has given rise to a growing recognition among linguists that events are not unanalyzable atoms but rather have an internal structure,

* I would like to thank Tsuguro Nakamura and two anonymous EL reviewers for helpful comments on earlier versions of this article. I am particularly grateful to one of the reviewers for very detailed comments and thought-provoking questions, which have substantially improved the content of the paper in many ways. Any remaining errors and shortcomings are, of course, my own.
organized around conceptual entities like action and change. Researchers have thus developed one model of event structure or another, particularly for the last two decades: some have employed logical tools, others algebraic notations, still others tree diagrams, and yet others image schemas, each with several variants. Little consensus has been reached, however, on what the proper representation of events should be like. The present study addresses this question, and argues for a new model of event representation called the “parallel event structure (PES).” Evidence is drawn primarily from some verbal and adverbial expressions in English, but it is suggested that essentially the same argument can be made on the basis of other languages’ corresponding expressions, too.

The argument proceeds as follows. Section 2 introduces the PES with a first illustration, featuring its novelty. Section 3 shows that the PES allows us to describe various types of events and verbs in a way that differentiates them from one another while capturing their commonalities. Section 4 shifts attention to arguments and modifiers, and discusses the role they play in the determination of event type, focusing particularly on telicity. It turns out through the discussion that an apparently disparate range of events can be analyzed uniformly and coherently in terms of the PES, which leads us toward a unified account of telicity in particular, and of event structure in general. Section 5 demonstrates that the PES is preferable to alternatives, pointing out that it has several advantages over them in terms of descriptive power and uniformity/coherence. Finally, section 6 summarizes the main points of the present study, and suggests a few directions for future research.

2. The Parallel Event Structure

The PES is a multi-layered model of event representation or such a structured event representation itself. It describes various types of events in terms of familiar conceptual entities like action and change—there is nothing novel about this. What is unique is their configuration: they are arranged on the vertical dimension (whereas they are arranged horizontally in most previous studies, which we shall see in section 5). This model of event representation leaves the horizontal dimension for time, over which the event unfolds.

Let us consider the following example, which provides a good first illustration:

(1) John closed the door.

This sentence denotes at least two events. One is a continuous action-
change event, which corresponds to Shibatani’s (1973/1975) “controlled
causation,” McCawley’s (1976) “continuous causation,” or Talmy’s (1985)
“extended causation,” where the actor continuously exerts his energy
against the patient (say, keeps pushing it) while it undergoes a change of
state. Another is an initiative action-change event (or so-called “ballistic or
onset causation”), where the actor only sets the patient in motion and stays
put. These two events can be described by the PES in a way that clearly
delineates them from each other while capturing their commonalities. 1

(1) John closed the door.

a. Continuous Action-Change Event

b. Initiative Action-Change Event

1 For the sake of simplicity, let us confine the discussion to cases where the door is
fully open before John’s action against it, although there are some other conceivable situ-
ations for which the sentence in (1) could be used felicitously (say, situations where the
door is half open or ajar before the action) and the PES can accommodate them as well
by relativizing the inchoative point (the point at which the event begins) to context.
Both structures have two major force-dynamic elements: ACTION (the exertion of energy by an entity) and CHANGE (an entity’s change of state in the broad sense that it covers change in any domain (e.g. shape, color, location, posture, etc.) and appearance/disappearance). They are referred to as “tiers” rather than more neutrally as, say, “components,” because the PES represents force-dynamic information on the vertical dimension and thus they are aligned vertically. The action tier has the conceptual function ACT(-ON) as its nucleus, which takes one or two arguments: the primary argument is ACTOR (an entity that exerts its own energy, typically a volitional human), and the secondary argument is ACTEE (an entity that receives that energy). They are also arranged vertically as subtiers within the action tier. The actor in (1) is JOHN, and the actee DOOR. They are thus put on the proper subtier in the above structures: JOHN on the actor tier, and DOOR on the actee tier (for the sake of simplicity, the determiner is left out). The pair of square brackets on the action tier thus represents the subevent in which the actor JOHN exerts his own energy against the actee DOOR, where the pair of double angle brackets around JOHN indicates that the actor is the external argument of the verb close, which is realized as the subject. The Greek subscripts on ACTEE and UNDERGOER (an entity that undergoes a change of state) are meant for coreference.

The temporal evolution of the action event is plotted on the horizontal axis. Notice that the left pair of square brackets in (1a) is put above the point t_i on the time arrow at the bottom, where the subscript i is meant for “inchoative.” This configuration signifies that the action event begins at that inchoative point. The action tier in (1a) has a solid line extending rightward from that square bracket pair, which indicates that the activity continues through time, and its termination is shown by the right pair of square brackets above the terminative point t_t and the absence of line to its right. The temporal phase between the inchoative and terminative points during which the action is carried out is called the “active phase.” On the other hand, the structure in (1b) does not have it, which signifies that the action is conceptualized as an instantaneous one, where t_{i/t} indicates that the

---

2 There are some cases where they are not coreferential, a few of which will be taken up later.

3 To be precise, there are cases where a preparatory action begins earlier than that point, and where it should be represented as such in the event structure, since it brings about some empirical consequences; see Miyakoshi (2009a) for discussion and illustration of them.
inchoative and terminative points on the action tier practically coincide.\footnote{To be precise, an apparently instantaneous action does have an active phase, and it can be highlighted in a certain context (see footnote 8).}

The change tier is also composed of two subtiers: UNDERGOER and STATE. The change event, if telic as in (1), is divided into three phases by the two boundary points (inchoative and terminative points). Let us call them “initial, mutative and resultative phases”: the initial phase is before change, the mutative phase in the middle of change, and the resultative phase after change. In the case of (1), the undergoer DOOR is open in the initial phase, and is closed in the resultative phase. Under the assumption made earlier (footnote 1), it usually takes some time to accomplish the door-closing event, and thus we can recognize the mutative phase in which the door undergoes a gradual change of state.

The change tier in (1) represents this transition of state as progress along a scale that is constituted of degrees of the gradable property CLOSED. Specifically, the undergoer does not even possess the minimal degree of that property (i.e., DOOR is fully open) at the starting point $t_0$, which is represented by the leftmost pair of square brackets on that tier (where the pair of single angle brackets around DOOR indicates that it is the internal argument of the main verb, which is syntactically realized as the direct object in this case). The undergoer then moves along the scale over time, gaining greater degrees of closedness, which is shown by a sequence of square bracket pairs arranged rightward on the change tier, each associated with a value of the property in question (e.g. $\text{CLOSED}_{\text{MIN}}$, $\text{CLOSED}_{\text{MAX}}$, and so on), where MIN and MAX are meant for “minimal and maximal values,” respectively. It finally reaches the maximal degree (i.e., DOOR is completely closed) at a telic point $t_{\text{tel}}$, which is tentatively defined as a definite endpoint of a change, and the presence of which is a necessary and sufficient condition for telicity (see sections 3 and 4 for discussion of its proper characterization). The resultant state of the undergoer is signified by the rightmost pair of square brackets on the change tier. A short solid and dashed line to its right indicates that the resultant state of the undergoer holds for an unbounded period of time, even if the action is not carried out any more after the terminative point (once one closes a door, it usually remains closed even if the actor does not exert his/her energy against it any more). This apparently trivial information should be represented in the resultative phase, because there are some empirical phenomena that are sensi-
tive to it, which will be illustrated in section 3.2.

The bold-facing of the rightmost bracket pair signifies the point at which the entire event is judged to be done; notice that the door is not considered to be closed until the point where it has attained the maximal value of the gradable property in question. Interestingly, a door-opening event is usually (if not always) judged to be done at a different point on the scale, which will be shown in section 3.2. In any case, I hereafter call such a point the “perfective point” of the event. If a given event unfolds up to this point, then one can felicitously use a perfective verbal sentence to denote it. In the case under discussion, for example, one can felicitously say “John has closed the door” at the telic point or thereafter, but cannot at any earlier point.

Having seen the two tiers (ACTION and CHANGE) in order, we should not miss their relationship. The undergoer’s change is brought about by the actor’s action against it (recall that the actee and the undergoer are coreferential in the case at hand, which is indicated by the Greek indices). This causal relation is signified by the downward arrow(s) between the two tiers, with the relational specification EFFECT.5

The horizontal arrow at the bottom represents the flow of time, where time is event-internal aspect time, as opposed to event-external tense time, which is set aside in this paper. The thick line on the arrow signifies the temporal profile of a given event (the span of time during which its evolution is tracked for its identification), which in turn indicates that it is symbolized by (a phrase/clause headed by) a verb.6 In the case of (1), the bold line extends from the initial phase through the resultative phase, because the action-change event denoted by the verbal sentence *John closed the door* ranges across all the three phases, as sketched above. The dots on the time arrow signify points in time. For example, as stated earlier, the dot above

---

5 It is often the case that the action is in turn driven by the actor’s volition and intention of changing the actee’s state. To differentiate such a willful and deliberate action from an inadvertent and unintentional one, the two-tiered structures in (1) might need to be elaborated into three-tiered structures. This issue is not discussed in this paper, but Miyakoshi (2008) addresses it, pursuing the avenue suggested just above.

6 I assume with Langacker that nonverbal expressions, unlike verbal ones, are atemporal in nature and lack a temporal profile (cf. Langacker (1987: chapters 6 and 7)). The presence of a thick line on the time arrow thus indicates that the event structure describes the meaning of (a phrase/clause headed by) a verb, and not of any other non-verbal expression.
t₁ in (1a) represents the inchoative point of the door-closing event.⁷

Overall, then, the PESs in (1) represent three types of information on the two dimensions. Specifically, they represent (i) force-dynamic information (the exertion of energy and its effect) vertically, (ii) simple aspectual information (the temporal evolution of each subevent) horizontally, and (iii) complex aspectual information (the two subevents’ parallel unfolding over time) two-dimensionally. Of particular importance among them is the last point (the two-dimensional representation of the complex aspectual information), which is one of the salient features of the PES and gives it power to describe various events in a way that differentiates them from one another while capturing their commonalities in a uniform way. With this feature, for example, the PES can clearly delineate continuous and initiative action-change events (so-called “extended causation” and “onset causation”), while simultaneously capturing their commonalities, as shown in (1a) and (1b). Particularly, its descriptive power figures prominently in the continuous action-change event: it explicitly describes the temporal paral-

⁷ An anonymous reviewer perceptively points out that this notation is not consistent with the scale structure introduced above. Recall first that under the present analysis, a change event is decomposed into a set of component states, each of which is represented with a pair of square brackets and is anchored to a dot on the time arrow by a vertical dashed line. And note that such decomposed states are customarily considered to be the minimal components of the event. Thus, the dots to which they are anchored should also be represented as the minimal components of time. The problem is that they are not represented that way in (1); assuming a mathematical characterization of a solid line as a set of dots, it follows that the numbered dots (say, t₁ and t₂) on the solid line arrow in (1) are not exactly adjacent to one another, and thus do not represent contiguous points in time. The reviewer suggests that this problem can be solved by adopting a representation like the following, where time is represented as an ordered set of dots that are spaced out in accordance with component states:

```
(i)
```

I basically concur with the reviewer on this matter, but here arises a different problem: the notation in (i) only works for certain limited cases. To be sure, it dovetails nicely with the scale structure representation of change events, but not with the solid line representation of continuous action events. To take (1a) for example, it fits well with the structure on the change tier, but not with the one on the action tier, which rather goes better with the original notation. This paper has no resolution of this dilemma, and tentatively adopts the notation introduced in (1). But we should bear in mind that this notational problem is yet to be solved for the proper representation of events.
lelism and coextensiveness of the actor’s action against the actee/undergoer and its change of state. Furthermore, the scale structure introduced above helps coherently represent the homomorphic relationship between the undergoer’s gradual change of state and the temporal evolution of the entire event. That is to say, the PES can capture the parallel progress of three types of (sub)events over time and represent it along the temporal dimension: (a) the continuous action subevent, (b) the gradual change subevent, and (c) the entire action-change event.

Let us see, then, how this multi-tiered structure works for other types of events and verbal expressions that denote them.

3. Events and Verbs: Examining the Descriptive Power of the Parallel Event Structure

This section shows that the PES can describe various types of events in a uniform and coherent way, giving necessary elaborations and qualifications to the structures introduced in the preceding section. As is now widely recognized, events and verbs that schematically denote them fall into several semantically-identifiable classes. The descriptive power of an event structure, therefore, can be measured by checking how well it can differentiate them from one another while capturing their commonalities. This section focuses on certain types of events, and puts the PES to empirical tests with English verbal sentences that denote them.

3.1. Action Events and Verbs

Let us begin with action events (ones where an actor carries out an action by exerting his/her own energy), which are usually denoted by (sentences headed by) verbs like unergatives and verbs of contact (or Vendler’s (1967) “activities” and Smith’s (1997) “semelfactives”). Verbs of this type do not inherently encode any change, and thus only receive an action tier under the present analysis. For example, a simple sentence with an unergative verb like Mary danced has the following structure:

(2) Mary danced.

```
ACTION
  \{ ACTOR
  \{ MANNER
    DANCE
  \}
```

TIME (EVENT)
The act of dancing is characterized by the actor’s continuous action in a certain manner. The action tier therefore has a manner tier as well as an actor tier, and its temporal unboundedness is indicated by the solid and dashed line extending after the pair of square brackets along the time dimension.

If the action is instantaneous by its nature, then it is described by a structure with no line that follows a pair of brackets on the action tier. For example, an instantaneous action event like ‘Mary hit the drum (once)’ is described by the following structure, where \( t_{i/t} \) indicates that the inchoative and terminative points practically coincide:

\[
\text{(3) Mary hit the drum (once).}
\]

A semelfactive verb like *hit* is often used for a repetitive action event as well, and it can also be accommodated by elaborating on the above structure as follows:

\[
\text{(4) Mary hit the drum (many times).}
\]

Notice here that these two-dimensional event structures allow us to delineate action events that differ in instantaneity and iteration from one another, and to describe the meanings of the verbal expressions denoting them in a way that explicitly shows their commonalities in force-dynamics (the actor’s exertion of energy and its characteristic manner) and differences in

---

As mentioned in footnote 4, an apparently instantaneous action event can be construed as having an active phase in a certain context. For example, if one records a single drum-hitting action by Mary on video and sees the film in slow motion, then he/she can clearly recognize the active phase and felicitously say “Mary is hitting the drum”; see Langacker (2008: 156) for a similar example.
aspect (durativity/instantaneity and iteration). Note also that the event structures in (3) and (4) signify that the actee is just that: it does not undergo a change of state. This is in contrast to the actee in (1) ‘John closed the door,’ which does undergo a change of state as the effect of the action; that is, ‘the door’ in (1) is both an actee and an undergoer. Let us hereafter call such a double-role participant “patient,” thereby differentiating it from just an actee like ‘the drum’ in (3)/(4). Under this analysis, the two types of events are structurally distinguished from each other: action events like (3)/(4) have a single-tiered structure with a simple actee as the internal argument of the main verb, whereas action-change events like (1) have a double-tiered structure with a patient.\footnote{Certain types of action-change events receive a triple-tiered structure, which will be illustrated in section 4.3.}

3.2. Change Events and Verbs

The durative vs. instantaneous opposition is observed in change events as well, and the model of event representation introduced above can successfully apply to them, with minor adjustments. For example, consider the following two structures, which denote durative and instantaneous change-of-state events, respectively:

(5) The door closed.

(6) The vase broke.

In contrast to the structure denoting a gradual change-of-state event in (5), the structure in (6) apparently has no mutative phase (let us put aside a special context like the one mentioned in footnote 8), and the resultative phase
is adjacent to the initial phase, which signifies the practical instantaneity of the undergoer’s change of state. That is, the adjacency of the two phases on the spatial representation mirrors their temporal contiguity in the actual event. This type of description is only possible if one takes advantage of the homology between space and time, which in turn requires a model of event representation like the one being argued for, where the temporal evolution of an event is plotted along a scale—another key feature of the PES in which it departs from many (if not all) of the previous models (see section 5 for comparison with them).

A flash of light is similar to vase-breaking in the practical instantaneity of the undergoer’s change, but is different in the continuity of the resultant state. If a vase breaks, then it remains broken for an unbounded period of time, which is signified by a solid and dashed line after the telic point in (6). In the case of light flash, on the other hand, the undergoer gets back to the initial state right after change. This type of event, therefore, should be described as in (7), where there is no solid and dashed line in the resultant phase:

\[(7) \text{ The light flashed (once).} \]

This is still another example which illustrates the point that the PES can structurally differentiate one type of event from another while capturing their commonalities.

A repetitive flash event can also be described just by arranging the two types of bracket pairs (which denote “on” and “off”) one after the other along the temporal dimension as in (8), where the solid and dashed line is meant to show an unbounded repetition of a pair of states (“on” and “off”) rather than of a simple resultant state:

\[(8) \text{ The light flashed (many times).} \]
Elaborating on the scale structure representation introduced in section 2, we can explicitly describe several other types of change events, too. For example, consider a door-opening event. As touched upon earlier, its perfective point is not identical with that of door-closing; specifically, door-opening is (usually) judged to be done when the undergoer attains a minimal degree of the gradable property denoted by the verb, whereas door-closing requires the undergoer to reach a maximal value on the property scale in question, for it to be true. In this sense, *open* can be taken as a “minimal-value-oriented” predicate, and *close* as a “maximal-value-oriented” one. This difference can be captured by describing a door-opening event as in (9); compare it with the structure for a door-closing event in (5):

(9) The door opened.

Recall that the bold-facing of a bracket pair indicates the perfective point of

---

10 They roughly correspond to so-called “partial and total predicates,” respectively (cf. Yoon (1996), Rotstein and Winter (2004), Kennedy and McNally (2005)), but I use the above terms since the total/partial dichotomy is too coarse-grained to make a semantic classification of change events/predicates, which will be illustrated shortly.

11 An anonymous reviewer points out that *open* is not always used as a minimal-value-oriented predicate, citing the following data as counterexamples:

(i) Groping his way in the direction of this last sound, Lanyard found the small side door ajar. He opened it, and hesitated a moment, looking out as though questioning the weather.

(http://www.online-literature.com/joseph-vance/lone-wolf/4/)

(ii) … the door, which I’d left slightly ajar, mysteriously opened.

(http://www.perrybrass.com/Excerpt2.html)

In these contexts, the verb *open* is naturally interpreted as denoting a change-of-state event where the undergoer’s resultant state is wide open. That is, these examples indicate that *open* can be used as a non-minimal-value-oriented predicate as well. This is in itself an intriguing observation in that it challenges the maximal vs. minimal (or total vs. partial) opposition. But it does not necessarily deny the observation stated above: the minimal positive aperture of a door suffices for door-opening to be judged to be done, whereas the attainment of the maximal value (complete closedness) is required for door-closing to be true. I thus use *open* as an example of a minimal-value-oriented predicate in this paper, leaving its non-minimal-value-oriented usage as a residual problem for future research. I am grateful to the reviewer for pointing out this problem.
the event, and notice that the position of the bold-faced bracket pair is the lower end of the scale (excluding the negative value) in (9), whereas it is the upper end in (5). They thus structurally represent the semantic difference between *open* and *close* stated above: door-opening is (usually) judged to be done only with the minimal positive aperture of the door, whereas the door is required to be closed completely for door-closing to be true.¹²

Let us next consider a fruit-ripening event. An event of this type has several intriguing properties, some of which are shared by door-closing/opening events, but others not. They are similar in that they denote a closed scale (which has an upper bound, beyond which higher values do not exist), as opposed to an open scale (which has no such upper bound) (cf. Hay et al. (1999)). They are also similar in that the maximal value can be denoted by the corresponding adjective and an endpoint-oriented adverb like *completely*, as shown in (10):

(10) a. The fruit is completely ripe.
    b. The door is completely closed/open.

However, as Kearns (2007: 45–46) perceptively points out, verbs of the *ripen* type do not necessarily entail the attainment of the maximal degree of the property scale they denote, which is evidenced by the observation in (11a) that the accomplishment interpretation of the verbal sentence can be cancelled by *not completely A* without contradiction.

(11) a. The fruit ripened in five days, but it wasn’t completely ripe.
    b. #The door closed in five seconds, but it wasn’t completely closed.

In the case of *close*, on the other hand, such cancellation results in a contradiction, as shown in (11b). This contrast verifies that *close* is a maximal-value-oriented verb in the sense stated above, but *ripen* is not. Is it, then, classified into the minimal-value-oriented type as *open* is? Obviously, it isn’t; for fruit-ripening, unlike door-opening, is not judged to be done with the attainment of the minimal value on the property scale in question. That is, the perfective point of *ripen* is neither the maximal value nor the minimal value on the scale of ripeness; as Kearns claims, it is rather a “standard value” on that property scale, which is conventionally and/or contextually determined. Specifically, a fruit is judged to have ripened when it reached

---

¹² This slight difference between ‘open’ and ‘close’ is observed in other languages as well, and brings about some empirical consequences, one of which is discussed in Miyakoshi (2009a).
(the lower bound of) a “standard value” on the scale of ripeness, which is usually determined by a complex interplay of such factors as the undergoer’s size, color, and sweetness. It is still judged to be ripe as long as it is within the region between (the lower bound of) the standard value and the maximal value on that scale, but the attainment of the maximal value is not required for the accomplishment of the event; hence the implication of completeness can be cancelled without contradiction. A change event of this type, therefore, should receive a structure like the following, where the standard value is indicated by the subscript ST on the property RIPE (notice that it differs from the structures for door-closing and door-opening):

(12) The fruit ripened in five days, but it wasn’t completely ripe.

(\(=\)(11a))

This fine-grained distinction between apparently similar change events can also be made in terms of the PES.

Hawk-eyed readers may have already noticed, but the above discussion leads us to reconsider the proper characterization of telicity. Recall that a telic point was tentatively defined earlier as a definite endpoint of a change, and that its presence is considered to be a necessary and sufficient condition for the event to be telic. Given the presence of “minimal- and standard-value-oriented” verbs like *open* and *ripen*, however, that definition needs to be revised; for the telic points of the change events denoted by (9) and (12) are not a simple endpoint of a change. Taking into account the convention/context-dependent nature of telicity shown just above, let us redefine it as follows: an event is telic iff it has a telic point, which is a conventionally and/or contextually determined perfective point of a change. This definition can properly characterize telic events including (12) as such, while correctly excluding atelic ones like simple action events (ones with no change as the effect of action) described in the last section. We shall come back to this issue in section 5, after describing various types of action-change events (including the familiar ones taken up in the previous studies of telicity) in the next two sections (§3.3 and §4).
3.3. Action-Change Events and Verbs

The PES can accommodate a variety of action-change events (ones with both an action and a change) as well. First of all, as sketched in section 2, action-change events like (1) ‘John closed the door’ are described by the PES in a way that (i) clearly delineates continuous and initiative action-change events, (ii) explicitly represents the temporal parallelism and coextensiveness of the actor’s action against the patient and its change of state, and (iii) coherently represents the homomorphic relationship between the patient’s gradual change of state and the temporal evolution of the entire event.

And again, taking advantage of the space-time homology, the PES can make a distinction between action-change events that differ in the instantaneity of change, too. For example, the event denoted by John broke the vase has the structure in (13), where the resultative phase is adjacent to the initial phase, which signifies their temporal contiguity.

(13) John broke the vase.

We can thus delineate events of this type from ones of the types in (1a) and (1b) as well, while simultaneously capturing their commonalities: e.g., (i) they are all action-change events with a patient (hence the prototypical double-tiered structures as opposed to non-prototypical ones, which will be illustrated later), (ii) they are all the same in the continuity of the resultant state of the undergoer (compare them with events of the type that will be taken up just below), and (iii) (1b) and (13) are similar in the instantaneity of the action.

The multi-tiered representation also allows us to differentiate events like ‘John held Bill’s arm’ from the ones discussed just above ((1) and (13)). Verbs of holding as well as verbs of closing and breaking denote an action-change event, but they are unique in that they denote an event where
the actor could (if not have to) keep exerting his/her energy against the patient even after the perfective point. Specifically, the act of holding an arm is said to be done once the actor holds the patient’s arm, and it usually does not take time; that is, it is a kind of instantaneous action-change event. An interesting feature of it is that one could (and often does) keep the resultant state for a moment, and it requires his/her continuous energy exertion. This type of event can be described as follows:

(14) John held Bill’s arm.

The point to be noted here is the presence of a parenthesized line to the right of the bracket pair on the action tier, which signifies the potential continuous action after the perfective point (which in turn brings about the continuation of the patient’s resultant state), whereas no such continuous action is carried out in the cases of door-closing and vase-breaking; hence the absence of a parenthesized line in the resultative phase on the action tier in (1) and (13).

Incidentally, one might here suspect that the event denoted by John held Bill’s arm is just an action event with no change as its effect. But it isn’t; for it is not judged to be done if the patient’s state does not change as the effect of the action (from ‘the state of not being held by John’s hand’ to ‘the state of being held by it’). This interpretation is supported by the fact that one can describe the resultant state by saying “Bill’s arm is in the state of being held by John’s hand.” In the case of simple action events like ‘Mary hit the drum,’ on the other hand, one cannot usually make such a statement: “# The drum is in the state of being hit by Mary.”
4. The Role of Arguments and Modifiers: Toward a Unified Account of Telicity and Event Structure

The preceding sections showed that the PES can explicitly describe certain types of events (action, change, and action-change events, each with several subtypes), but the particular events taken up there are all simple in the sense that they are all denoted by a sentence (i) whose predicate is a single verb, (ii) whose argument is a count and singular noun phrase, and (iii) which has no adverbial modifiers—one that roughly corresponds to what Smith (1997: 54) calls a “maximally simple sentence.” That is, we have only seen cases where the aspectual type of a given event is uniquely determined by the main verb. This section deals primarily with familiar sets of events whose type varies depending on the presence/absence and the nature of the verb’s argument(s) and/or a certain type of adverbial modifier or secondary predicate. Particularly discussed are the questions of what role they play in the determination of telicity and how it is structurally described. It turns out through the discussion that an apparently disparate range of events can be analyzed uniformly and coherently in terms of the parallel event and scale structure proposed above, which leads us toward a unified account of telicity in particular and of event structure in general.

4.1. Events with a Physical Amount Scale

As many researchers have pointed out, the telicity of an event is determined compositionally by some factors, and the boundedness of the internal argument plays a particularly important role in many cases (cf. Verkuyl (1972, 1993), Dowty (1979, 1991), Jackendoff (1991, 1996), Krifka (1992, 1998), Tenny (1994), and many others). Consider the following stock examples:

(15) a. John ate an apple.
    b. John ate custard.
    c. John ate apples.

The event denoted by (15a) is telic, whereas the ones denoted by (15b/c) are not. This difference is attributed to the boundedness of the internal argument. The argument of this type (and its thematic role) has been called by

---

13 Sentences of the (15a) type can receive an atelic interpretation in certain contexts (cf. Smollet (2005: 49–50)), which lends further credence to the context-dependent nature of telicity mentioned earlier, but let us leave this issue aside here.
a variety of names in the literature (the most widely adopted among them being Dowty’s “incremental theme”), but I just call it a “patient” in the sense defined earlier, because it turns out to be essentially no different from the ordinary patient as an aspectual determinant. In any case, it is bounded in both quantity and number in (15a), which delimits or “measures out” the entire event, whereas the patients in (15b/c) are unbounded in quantity or number, which makes the event atelic. The PES can describe these events in a way that captures their difference in telicity and its origin: the (un-) boundedness of the patient:

(16) a. John ate an apple.

b. John ate custard.
c. John ate apples.\footnote{The sentence \textit{John ate apples} has both sequential and simultaneous event readings, but only a sequential event is described in (16c) since people usually eat two or more apples that way.}

The multi-tiered structure in (16a), just like the one in (1a), explicitly describes a number of force-dynamic and aspectual properties of the telic action-change event, which include the following: (i) the temporal parallelism and coextensiveness of the actor’s action against the patient and its change of state (specifically, the first bite of an apple by John brings about a slight change in its amount, which is quantified as the traversal of the physical amount scale from zero to one, the second bite makes it progress further from one to two, and so on), (ii) the homomorphic relationship between the patient’s gradual change of state and the temporal evolution of the entire event (the physical amount of the patient consumed is homomorphically mapped to the temporal progress of the apple-eating event), (iii) the telicity of the event (it is temporally bounded at the telic point), and (iv) the main determinant of its telicity (it is the patient’s boundedness in both physical quantity and number that gives the entire action-change event a definite endpoint). If the patient is not bounded in physical quantity or number, then the event lacks such an endpoint, which in turn makes it atelic, as shown in (16b) and (16c), respectively.

Note here that the present analysis of food-eating events departs from previous analyses in terms of the treatment of verbs of consumption. Previous studies have aspectually classified them as “activities” in the sense of Vendler (1967) (see Dowty (1979: 67) for example), but the present study takes them as action-change verbs rather than as simple action verbs, regardless of
the telicity of the events denoted by (sentences headed by) them. This is because they inherently encode the patient’s change of quantity/place as well as the actor’s action; for any eating/drinking event is not judged to be done unless a part or whole of the food/liquid undergoes a change of quantity/place (specifically, unless it is taken into the actor’s body through his/her mouth). It thus follows that verbs like eat/drink and events they denote, telic or atelic, should be analyzed as verbs/events that have a double-tiered structure like the above, rather than as ones that only have an action tier.15

4.2. Events with a Spatial Distance Scale

This analysis can be extended successfully to events with a spatial distance scale as well, regardless of the syntactic type of an element that delimits the event. Let us first consider events with a spatial path or “holistic theme,” which are denoted by standard examples like the following (cf. Dowty (1991: 569), Tenny (1994: 17–18)):

(17) a. Bill climbed the ladder.
   b. Sue walked the Appalachian Trail.

Sentences of this type have a direct object that refers to a path of a definite length, along which the actor/undergoer moves.16 They have both telic and atelic readings, but the telic reading is attributed to the boundedness of the path or “holistic theme” in spatial distance; for example, the length of the ladder measures out the climbing event in (17a). This type of event can be described as in (18), where it is assumed for the sake of discussion that Bill was at the lowest point of the ladder when he began to climb it:

---

15 An action of the eating/drinking type, in fact, can be construed as bringing about a change of the actor as well as of the actee, and a description of such an event requires a triple-tiered structure that consists of one action tier and two change tiers, which will be illustrated in section 4.3.

16 Bill/Sue in (17) should be taken not just as an undergoer but rather as an actor as well, since he/she undergoes a change of location by exerting his/her own energy. This characterization is independently needed for the proper description of many linguistic phenomena; for example, it is required to draw a line between the following pair of sentences:

(i) Laura deliberately rolled the course.
(ii) *The package rolled the course. (Tenny (1995: 218–19))

Laura in (i) is an actor as well as an undergoer, and thus qualifies as the external argument of the verb roll, which in turn leaves the object position for a path argument; hence the sentence is acceptable. On the other hand, the package in (ii) is just an undergoer (hence the internal argument of roll), and thus cannot occur as the subject of a transitive sentence with a path argument as the object.
(18) Bill climbed the ladder.

This event is similar to the one in (16a) in that they are both an action-change event which is telic and whose delimiter is syntactically realized as the direct object. On the other hand, they differ in the type of the delimiter: it is the patient that undergoes a change of state in (16a), whereas it is the path through which the actor undergoes a change of position in (18). They differ in the type of scale, too: it is a physical amount scale in the former, while it is a spatial distance scale in the latter. Overall, however, they receive a unified and coherent account of telicity and structural description of events under the PES analysis.

4.3. Events with an Optional Delimiter

Essentially the same account can be made for many other events as well, including ones that are delimited by an optional modifier rather than by an obligatory argument. Consider the following sentence (still another stock example), which denotes a telic action-change event with a spatial distance scale, just as the sentences in (17) do:

(19) Bill pushed the cart to the store.

A sentence of this kind differs from one like (17) in the type of delimiter: it is a modifier PP in (19) whereas it is an argument NP in (17). This type of sentence can be described as in (20), where an inherently unbounded action-change event is delimited by the end of a path that is denoted by the directional PP (“ST” on the place tier in (20) is meant for “STORE”):\(^{17}\)

\(^{17}\) To be precise, the PES in (20) is not descriptively sufficient in that it only represents one of the two distinct actions performed simultaneously by the actor in this event: Bill’s action against the cart in a certain manner is explicitly represented on the action tier, but
Here again, the telicity of the event is attributed to the boundedness of the path along which the actor/patient undergoes a change of location. It thus follows that if the PP is replaced by one that denotes an unbounded directional path, then the event becomes atelic, which is described in terms of the PES as follows:

(21) Bill pushed the cart toward the store.

When it comes to resultative constructions (RCs), an adverbial (and predicative) AP as well as PP can delimit an inherently unbounded event. For
example, consider the following pair of sentences:

\[(22)\]  

a. John pushed the door.  
b. John pushed the door open.

The sentence in (22a) denotes an atelic action event, which can in principle go on for an indefinite period of time, and does not necessarily bring about the patient’s change of state. Thus, it should be described by a single-tiered structure like the following, where the pairs of parentheses in the resultative phase indicate that the act of door-pushing may be carried out continuously, though it does not have to be for the sentence to be true:

\[(23)\]  

John pushed the door.

The addition of the AP *open*, however, shifts it to a telic action-change event, which is delimited by the resultant state of the patient it denotes. That is, the optional resultative phrase (RP) functions as the delimiter of an event—yet another well-known observation. Under the present analysis, this type of RC receives a double-tiered structure like the following:\[19\]

\[\text{RCs of this type, however, are extremely restricted in distribution, and RCs with a nominal resultative phrase are mostly unacceptable, as shown in (iii) and (iv):}\]

(iii) He shot the robber {dead/to death/*a corpse}.  
     (Simpson (1983: 153, note 2))
(iv) She grounded the coffee beans {powdery/to a fine powder/*a fine powder}.  
     (Carrier and Randall (1989: 45))

Moreover, the apparently nominal resultative phrases in (i) and (ii) serve the function of denoting a property rather than referring to a thing. They should be analyzed, therefore, as an element whose semantic head is the adjective ‘yellow/red,’ perhaps with a reanalysis of the preceding part of the apparent NP ‘a pale shade of/a revolting shade of’ as a kind of degree modifier. In any case, RCs of this type are set aside in the discussion to follow, and the grammatical status of the resultative phrase is also left untouched in this paper; see Miyakoshi (2007, 2009b) for discussion of it.

\[19\] The sentence in (22b), just like the one in (1), denotes both continuous and initiative action-change events, but for the sake of simplicity, only the latter is described in (24).
Recall here that *open* is (basically) a minimal-value-oriented predicate, and thus the event in question is judged to be done with the minimal positive aperture of the door as the effect of John’s action against it (though it may and often does go on moving toward the maximal value point of the scale). This type of RC, therefore, constitutes yet another piece of evidence for the claim made earlier that the telic point is not necessarily a simple endpoint of a change, as well as for the claim that the addition of an optional RP brings about a shift in telicity.

This does not necessarily mean, however, that an RP *always* changes the telicity of a given sentence. For example, the AP *black* and the PP *into pieces* in the following (b) sentences are both considered to be RPs, but they do not change the telicity of the event:

(25)  
(a) Mary dyed her hair.  
(b) Mary dyed her hair black.

(26)  
(a) John broke the vase.  
(b) John broke the vase into pieces.

This is because the (a) sentences denote a telic event independently of the RP, and thus it only serves the function of making an additional specification of the resultant state of the patient. RCs of this type therefore receive structures like the following, where the patient’s resultant state is doubly specified by the predicative constants of the main verb and of the RP.20

20 The structures in (27) and (28) describe prototypical hair-dyeing and vase-breaking events, respectively: the former is a continuous action-change event in the sense stated earlier, whereas the latter is an instantaneous event in terms of both action and change.
We have thus far seen two familiar types of RCs that differ in the aspec-tual function of the RP: one like (24) where the addition of an RP brings about a shift in telicity along with the addition of a change tier, and one of the (27)/(28) type where it does not give rise to such changes. Sentences like the following illustrate a third type of RC, which is, unlike the above two, nontrivial and requires further elaboration of the PES to describe:

(29) a. Mary arranged the leeks flat on the plate.
b. The students lined up straight in the schoolyard.

Sentences of this type intuitively seem similar to RCs of the above types, but are different in that the post-verbal AP, unlike the above ones, does not describe the resultant state of the patient. For example, as Kageyama (1996: 236) points out, arranging leeks on a plate does not necessarily bring about their becoming flat. It seems, however, that flat in (29a) does denote a resultant state of something. What is it, then? I argue that it is an aggregate of leeks, which is created by the action of arranging individual leeks
on a plate. That is to say, the AP flat denotes the state of a product (rather than a patient) of the action-change event. The product by definition comes into existence in the resultative phase, and it is this product-orientedness of the AP that gives us the intuition that the sentence in question seems a kind of resultative.\(^{21}\) Elaborating on the PES further, we can describe this type of sentence as follows:

\[(30)\] Mary arranged the leeks flat on the plate.

This is a triple-tiered structure, which consists of one action tier and two change tiers. The first change tier (CHANGE\(_1\)) encodes a change of place/state that the patient (individual leeks) underwent, and the second change tier (CHANGE\(_2\)) denotes the emergence of the product (an aggregate of leeks) and its state.\(^{22}\) The correspondence between the patient and the product is indicated by the Greek subscripts \(\alpha/\alpha'\) with pairs of angle brackets. What should be particularly noted here is that this elaborated PES explicitly represents not only (i) the thematic relations that hold among the participants of the event denoted by the main verb, but also (ii) the predication relation that holds between the AP flat and its implicit predicant ‘an aggregate of leeks,’ (iii) its resultative nature, (iv) the correspondence between

\(^{21}\) Whether or not it should be regarded as a kind of RC is not uncontroversial; see Miyakoshi (2007, 2009b) for detailed discussion of this issue.

\(^{22}\) Recall that the present paper uses the term CHANGE in a broad sense, including emergence as well as a change of state, location, posture, etc.
that implicit argument of the secondary predicate and the explicit argument of the verb, and here again (v) the parallel unfolding of multiple subevents over time. The sentence in (29b) can also be described in essentially the same way: the crux is that what *straight* denotes is the state not of individual students but rather of a group of students who lined up in the schoolyard. That is, it is also a product-oriented RP.

There also exist sentences that denote a telic action-change event whose optional delimiter is an *actor*-oriented RP, which can be taken as RCs of a fourth type—another nontrivial phenomenon that deserves explanation. As many researchers have pointed out, sentences with a subject-oriented secondary predicate like the following are not acceptable as an RC in English, perhaps due to the controversial “Direct Object Constraint” (cf. Levin and Rappaport Hovav (1995: chapter 2), Miyakoshi (2006), and the references cited therein):

(31) *John ate the steak full. (on a resultative reading)

The corresponding sentence in Japanese, however, is completely acceptable on a resultative reading:

(32) Itiroo-ga suteeki-o onakaippaini tabe-ta.

Ichiro-NOM steak-ACC full eat-PAST

‘Ichiro ate the steak and became full as a result.’

There are some reasons why this sentence with a subject-oriented RP is acceptable in Japanese (whereas the English counterpart is not), one of which is that the actor of the food-eating event can be construed as an undergoer as well, because the eater’s internal state changes (specifically, he/she becomes less hungry) as his/her action proceeds.\(^{23}\) Thus, this type of event also receives a triple-tiered structure that consists of one action tier and two change tiers, though this time the second change tier is for an actor/undergoer rather than for a product; compare (33) with (30):\(^{24}\)

\[^{23}\text{See Miyakoshi (2006, 2007, 2009b) for other reasons.}\]

\[^{24}\text{In this case, to be precise, the patient denoted by the direct object does not have to be bounded in quantity and number. This is because the main delimiter of this sentence is the RP and it is predicated of the subject.}\]
In any case, these analyses of non-standard RCs evince the strong descriptive power of the PES: it can even describe complex events like (30) and (33) where two change subevents over and above an action subevent unfold in parallel.

5. Advantages of the PES over Alternatives

The preceding sections showed that the PES can describe various types of events explicitly, capturing their commonalities and differences. This section compares the PES with some alternative models of event representation proposed by previous studies, pointing out that it has several advantages over them in terms of descriptive power and uniformity/coherence.

First of all, recall that the PES represents (at least) three types of information on the two dimensions: it mainly represents (a) force-dynamic information vertically, (b) simple aspectual information horizontally, and (c) complex aspectual information two-dimensionally. More specifically, just one PES like (1a), repeated here as (34), represents all of the information listed below:
(34) John closed the door. (a continuous action-change event reading)

(35) a. force-dynamic information:
   i. the actor’s action against the patient
   ii. the patient’s change of state
   iii. the patient’s change of state was the effect of the actor’s action against it

b. simple aspectual information:
   i. the act of door-closing began at the inchoative point
   ii. it continued until the terminative point
   iii. no action was carried out after that point
   iv. the door was fully open in the initial phase
   v. it underwent a gradual change of state in the mutative phase
   vi. it was completely closed in the resultative phase

c. complex aspectual information:
   i. the patient’s change of state began with the actor’s action against it
   ii. the action and change events unfolded in parallel over time
   iii. the patient’s change of state ended with the end of the actor’s action
   iv. the patient’s change of state had a homomorphic relationship with the temporal evolution of the entire event

Second, the PES, if properly elaborated and qualified, can clearly delineate various types of events from one another while capturing their commonalities in terms of force-dynamic and aspectual parameters, some of which are orthogonal to each other. In terms of force-dynamics, specifically, events are classified into several types, three of which are focused on
in this paper: (a) action, (b) change, and (c) action-change events. Events of each type are cross-classified into subtypes by an orthogonal aspectual parameter: durativity/instantaneity. Thus, action events are subcategorized into durative and instantaneous types (e.g. dancing in (2) vs. drum-hitting in (3)). Change events also subdivide into two types in terms of this parameter (e.g. door-closing in (5) vs. vase-breaking in (6)). When it comes to action-change events, a durative vs. instantaneous distinction is made at three levels, and fine-grained classifications are accordingly made: (i) only at the active phase on the action tier (while the mutative phases are both durative) (e.g. continuous vs. initiative door-closing by an actor in (1a) and (1b)), (ii) at the active and mutative phases in parallel on the action and change tiers (e.g. continuous door-closing vs. instantaneous vase-breaking by an actor in (1a) and (13)), and (iii) at the active and resultative phases in parallel on those two tiers (e.g. vase-breaking vs. arm-holding by an actor in (13) and (14)). These force-dynamic and aspectual classifications of events are summarized in (36):

(36) The durative vs. instantaneous distinction is made at the levels of:
   a. only the active phase on the action tier (with no change event) (e.g. (2) vs. (3))
   b. only the mutative phase on the change tier (with no action event) (e.g. (5) vs. (6))
   c. i. only the active phase on the action tier (with a durative mutative phase) (e.g. (1a) vs. (1b))
      ii. the active and mutative phases in parallel on the action and change tiers (e.g. (1a) vs. (13))
      iii. the active and resultative phases in parallel on those two tiers (e.g. (13) vs. (14))

Third, the PES can make fine-grained classifications of events in terms of telicity, too. We have seen that an apparently disparate range of events can be broadly classified into two types with respect to the main determinant of their telicity:25 (a) events whose telicity is internally determined by the main verb, and (b) ones whose main telicity determinant is a verb-external element like the internal argument or a certain type of modifier. Events

25 For the sake of simplicity, I here confine the discussion to single events, thereby excluding cases where the same event is iterated (e.g. multiple drum-hitting and apple-eating in (4) and (16c)).
of the first type, which were dealt with in sections 2 and 3, can be classified further into three subtypes, depending on the orientedness of their telic point: (i) maximal-value-oriented (e.g. door-closing in (5)), (ii) minimal-value-oriented (e.g. door-opening in (9)), and (iii) standard-value-oriented (e.g. fruit-ripening in (12)). On the other hand, events of the second type, which were mainly taken up in section 4, can also be subcategorized into several types, depending on two parameters: (i) the nature of the scale they denote ((α) a property scale (e.g. door-opening in (24)), (β) a physical amount scale (e.g. apple-eating in (16)), and (γ) a spatial distance scale (e.g. ladder-climbing in (18) and cart-pushing in (20))), and (ii) the nature of the element that the optional delimiter is predicated of ((α) a “basic” patient, one that is inherently specified as such by the main verb (e.g. hair-dyeing in (27) and vase-breaking in (28)), (β) a “derivative” patient, one that is not lexically specified but is rather contextually construed as such (e.g. door-opening in (24)), (γ) a product (e.g. leek-arranging in (30)), and (δ) an actor/undergoer (e.g. steak-eating in (33))). This is summarized as follows:

(37) classifications of events in terms of the main telicity determinant:

a. verb-externally-determined telic events:
   i. maximal-value-oriented (e.g. (5))
   ii. minimal-value-oriented (e.g. (9))
   iii. standard-value-oriented (e.g. (12))

b. verb-externally-determined telic events:
   i. the nature of a scale:
      α. a property scale (e.g. (24))
      β. a physical amount scale (e.g. (16))
      γ. a spatial distance scale (e.g. (18) and (20))
   ii. the nature of the element that the optional delimiter is predicated of:
      α. a basic patient (e.g. (27) and (28))
      β. a derivative patient (e.g. (24))
      γ. a product (e.g. (30))
      δ. an actor/undergoer (e.g. (33))

Despite this disparate range of events that differ from one another in several parameters, the PES can capture their commonalities in telicity and give them a uniform and coherent structural description. To be sure, the events taken up in the preceding sections are far from exhaustive in terms of type and number, but the analysis of various events presented above at least suggests a way toward a unified account of telicity in particular, and of event structure in general.
Let us next see if and to what extent other models of event structure can accommodate pieces of information of the (35) type, and can make fine-grained classifications of events while capturing their commonalities, as done in (36) and (37).

As touched upon at the beginning of this paper, a large body of research has been done on event structure over the past forty years, and particularly the last two decades have seen the development of structured representations of events, to which the present study owes much. Previously proposed event structures can be classified into four groups, depending on how much force-dynamic and aspectual information they capture, and on how they do it. In what follows, I examine them in order, and point out that they have limitations in terms of descriptive power and uniformity/coherence, although they capture certain facets of events in one insightful way or another.

A first group is characterized as structures that represent events linearly, in the sense that both force-dynamic and aspectual information is represented on one dimension. For example, Langacker’s (1991) billiard-ball model and the causal/action chains advocated by Croft (1991) and Kageyama (2001) are of this type:

(38) a. Langacker’s (1991) Billiard-Ball Model

\[ \text{S} \xrightarrow{\text{X}} \text{O} \]

*The thief opened the window with a crowbar.*

\[ \text{S} \xrightarrow{='='} \text{O} \]

*A crowbar opened the window.*

\[ \text{S} \xrightarrow{='='} \]

*The window opened.*

(b. Croft’s (1991) Causal Chain

John hand hammer boulder (boulder) (boulder)

\[ \cdot \rightarrow \cdot \rightarrow \cdot \rightarrow \cdot \rightarrow \cdot \rightarrow \cdot \rightarrow \cdot \]

VOL Grasp Contact Change State Result State

*John broke the boulder with a hammer.*
Notice that these structures all capture force-dynamic information diagrammatically, and share one characteristic feature: they all arrange action and change on the horizontal dimension, which putatively represents the temporal evolution of the event as well. In these structures, consequently, any action-change event is represented in such a way that the change begins after the action is over. Event structures of this type might barely do for instantaneous action-change events like boulder-breaking and watermelon-cutting (which is presumably why the proponents of this one-dimensional model have always employed such examples as (38) for the purpose of illustrating their point). But they fail to properly describe action-change events of the (1a)/(34) type, because they cannot in principle capture the subevents’ parallel unfolding over time, at least as they now stand. That is, one of the critical problems with event structures of the (38) type is that they fail to accommodate complex aspectual information of the (35c) type. Moreover, they do not make most of the event classifications summarized in (36) and (37), either. The PES thus stands apart from them in these respects.

A second group is one that captures aspectual information to some extent, but not sufficiently. What belongs to this class includes the event structures proposed by Pinker (1989), Jackendoff (1990), Pustejovsky (1995), and Kageyama (1996):

(39) a. Pinker’s (1989) Structure

\[
\text{EVENT} \quad \text{ACT} \quad \text{THING} \quad \text{THING} \quad \text{MANNER} \quad \\text{“eating”} \quad \rightarrow
\]


\[
\begin{align*}
\text{<Action>} & \rightarrow \text{<Change>} \rightarrow \text{<Result State>}
\end{align*}
\]

father acts on watermelon with a knife

watermelon undergoes a change of state

watermelon is cut in half

*The father cut the watermelon in half with a knife.* (p. 6)
b. Jackendoff’s (1990) Structure
   i. *Bill dragged the car down the road.* (entraining)
      
      \[
      \begin{align*}
      \text{CS}_{\text{entrain}}^+ & \left( \{\text{BILL}, \text{GO} ([\text{CAR}], [\text{DOWN} [\text{ROAD}]])\} \right) \\
      \text{AFF}^- & \left( \{\text{BILL}, \text{CAR}\} \right)
      \end{align*}
      \]

   ii. *Bill threw the ball into the field.* (launching)
      
      \[
      \begin{align*}
      \text{CS}_{\text{launch}}^+ & \left( \{\text{BILL}, \text{GO} ([\text{BALL}], \left[ \text{FROM} [\text{BILL}] \rightarrow [\text{IN} [\text{FIELD}]] \right] \} \right) \\
      \text{AFF}^- & \left( \{\text{BILL}, \text{BALL}\} \right)
      \end{align*}
      \]
      (p. 138)

c. Pustejovsky’s (1995) Structure
   i. $e_{<_{oc}}$
   
      \[
      \begin{align*}
      e_1 & \quad e_2 \\
      \text{EVENTSTR} & \left[ \begin{array}{c}
      \text{build} \\
      E_1 = \text{PROCESS} \\
      E_2 = \text{STATE} \\
      \text{RESTR} = <_{oc}
      \end{array} \right]
      \end{align*}
      \]

   ii. $e_{o_{oc}}$
   
      \[
      \begin{align*}
      e_1 & \quad e_2 \\
      \text{EVENTSTR} & \left[ \begin{array}{c}
      \text{accompany} \\
      E_1 = \tau_i \\
      E_2 = \tau_i \\
      \text{RESTR} = o_{oc}
      \end{array} \right]
      \end{align*}
      \]

   (pp. 69–71)

d. Kageyama’s (1996) Structure

   Event (Accomplishment)
   
   \[
   \begin{align*}
   \text{Upper Event} & \quad \text{CONTROL} & \quad \text{Lower Event} \\
   \text{ACT} & \quad \text{BECOME} & \quad \text{STATE} \\
      \end{align*}
      \]
      (p. 90)
These structures are more elaborated than the above ones, and their descriptive power is accordingly enhanced. Particularly, Jackendoff (1990) and Pustejovsky (1995) are aware of the difference between continuous and initiative action-change events, and attempt to capture it in one way or another: specifically, with the verbal qualifications “entrain vs. launch” for CS(CAUSE) in Jackendoff (1990), and with the symbolic annotations “○ vs. <” in Pustejovsky (1995), where the circle and the inequality sign indicate respectively that the two subevents temporally overlap and are strictly sequential.26

These structures, however, are not sufficient for a full description of events, because several pieces of critical information remain unspecified. For example, there is no part in Jackendoff’s structure that exclusively represents the patient’s state in the resultative phase, although it has a part that represents its change of state (the pair of square brackets headed by the GO function). Conversely, there is no part in Pustejovsky’s structure that represents the patient’s change of state, although it has a part that represents its resultant state ($e_2$ represents just a state, and $e_1$ only encodes an action process). Kageyama’s structure neatly remedies these problems by positing separate representations for the change of state (BECOME) and for the resultant state (STATE/BE). It is, however, not elaborated in a way that can draw a line between continuous and initiative action-change events.

And again, this causative delineation is not the only distinction that needs to be made, but is rather just one instance of the more general durative vs. instantaneous distinction discussed earlier (recall (36)). Previously proposed structures like the above, however, mostly fail to make that distinction; nor for that matter do they make a fine-grained phase demarcation in the first place. In cases where a durative/instantaneous distinction appears to be made, arbitrary conventions tend to be crucially used; for example, there is no principled reason for assuming that the circle and the inequality sign represent continuous and initiative action-change events, respectively. Moreover, the above structures do not represent events in a way that can classify them in terms of telicity, either. They thus fail to capture the commonalities and differences in telicity between various types of events as the PES does (recall (37)).

Scale structures proposed by current studies (Hay et al. (1999), Kennedy

---

26 To be precise, however, the illustration in (39ci) is not appropriate because the act of building something and its coming into existence are not strictly sequential.
and McNally (2005), Kearns (2007), and many others) constitute a third group. They have focused particularly on aspectual issues associated with gradable predicates, and have developed structures like the following:

\[ \exists e, d[\text{INCREASE}(\text{long(rope)})(d)(e)] \quad \text{Kim lengthened the rope.} \]

(b. Kearns’s (2007) Scale Structure
\[ x \text{ verbed.} \]
\[ d' = \text{lower bound of standard value for } A \text{ in context} \]

There is no doubt that they have also made substantial contributions to the elucidation of (a certain part of) event structure, and the present study is based partly on their insights as well. It is, however, unclear how their scale structure fits into the overall structure of action-change events. Scale structures like (40) may do for change events denoted by intransitive verbs/sentences, but are yet to be elaborated for (causative) transitives, with the addition of a (sub)part that represents force-dynamic information. Researchers pursuing this approach recognize this necessity and intentionally leave that part out of the structure (see Hay et al. (1999: 132, 142), Kearns (2007: 38–39, 44)), but it would not be so easy to overcome this residual problem as they seem to expect. To take (40a) for example, they need to develop at least two structures for events denoted by Kim lengthened the rope (continuous and initiative action-change events) in a way comparable to (1a) and (1b), but it seems unclear how they do it under their framework. In any case, it would be safe to claim that scale structures like (40), as they stand, have limitations in descriptive power (specifically, in that they can only accommodate some of the properties in (35)–(37)).

A fourth group goes a step further, developing two-dimensional representations like the following:

\[
\begin{align*}
&\begin{bmatrix} [1d]^\alpha \end{bmatrix} \\
&\begin{bmatrix} [1d]^\alpha \end{bmatrix} \\
&\begin{bmatrix} \text{Sit PERFORM ([Thing X], [Thing 0d]); [Time 0d]} \end{bmatrix}
\end{align*}
\]


\[
\text{John closed the door.}
\]

Details aside, these elaborated event structures crucially differ from other alternatives in that they not only describe both force-dynamic and aspectual information, but also represent them on different dimensions. This innovative departure, coupled with a set of formal operations, gives them strong descriptive power so that the pieces of information enumerated in (35) could mostly be accommodated.

There are, however, inherent limitations with them. One is that the flow of time is represented on the vertical dimension: both Jackendoff and Iwamoto take the temporal evolution of an event as a one-dimensional projection of a state, and represent it upward. Another is that the vertical axis is not represented as a scale. It seems that these limitations jointly
make their representation unnecessarily complicated, which in turn makes it difficult to describe crucial information that could otherwise be captured more explicitly. As is well-known, the temporal evolution of an event is customarily plotted along a scale on the horizontal axis in many fields of study. This is presumably because it is the most explicit way to represent an event on a two-dimensional medium like a sheet of paper or a screen. It is this scalar representation that allows us to take advantage of the space-time homology for the purposes of describing certain facets of events; for example, recall that it was required as a prerequisite for the descriptive delineation between durative and instantaneous events like door-closing in (5) and vase-breaking in (6). On the other hand, feature-based representations like (41) do not enjoy this advantage. I know of no reason, empirical or conceptual, not to employ this convenient and explicit mode of representation for the description of time flow, and doubt that linguistics should be radically different from other disciplines in this respect.

And again, event structures like (41) are not elaborated in a way that can make such fine-grained classifications of events as the PES does, either. They thus fail to capture many, if not all, of the observations in (36) and (37). This is partly because they decompose events neither into tiers nor into phases, and also because they classify events neither in terms of the orientedness of their telic point nor in terms of the nature of the element that the optional delimiter is predicated of.

In short, the previous studies examined in this section are all insightful and can each accommodate some of the properties of events listed in (35)–(37), but not all. One might here be tempted to pursue a “hybrid” model in which force-dynamic information is captured by, say, image schemas, and aspectual information by, say, scale structures. However, it is unclear how such entirely distinct structures are unified into one coherent event structure; recall that both force-dynamic and aspectual information needs to be represented coherently to capture the properties of various events, particularly of action-change events like (1) and (13)–(33). I am aware that the meanings of linguistic expressions have many facets, and that several different modes of semantic representation are likely to be required for their proper description. But at least for the description of events of the types dealt with in this paper, I argue, the PES is preferable to alternatives in terms of descriptive power and uniformity/coherence, and the pieces of evidence presented above jointly corroborate it.
6. Concluding Remarks: Summary and Directions for Future Research

This article has argued for the PES, adducing several pieces of evidence primarily from verbal and adverbial expressions in English. One of the most characteristic features of the PES is the multi-tiered representation of a multi-faceted event: it mainly represents (i) force-dynamic information on the vertical dimension, (ii) simple aspectual information on the horizontal dimension, and (iii) complex aspectual information on the two dimensions. Another key feature of the PES is the parallel scalar representation of (a) the flow of time and (b) the unfolding of an event (and its subevent(s): the actor’s action and/or the undergoer’s change of state). These two features jointly differentiate the PES from alternatives like (38)–(41) in several respects: (α) the number of dimensions employed for event representation (two rather than one; cf. (38) and (40)), (β) the dimension for the representation of force-dynamic information (vertical rather than horizontal; cf. (38) and (39)), (γ) the dimension for the representation of aspectual information (horizontal rather than vertical; cf. (41)), and (δ) the nature of that aspectual representation (scalar rather than feature-based; cf. (41)). This multi-tiered and scalar structure gives each event an explicit and detailed description (recall (34) and (35)). It can also delineate various types of events from one another clearly, while capturing their commonalities in a uniform and coherent way (recall (36) and (37)). It should therefore be favored over alternatives in terms of descriptive power and uniformity/coherence, which in turn gives justification for the two-dimensional enrichment of event structure that has been argued for throughout the paper.

This article, on the other hand, still leaves many questions unanswered. One is how the PES applies to events of the types that are not dealt with here, which include (i) simple state events denoted by verbs like be and resemble, (ii) action-state events like ‘keeping the meat fresh’ and ‘staying home,’ and (iii) non-agentive causative events denoted by sentences like The alarm clock awakened John and The news surprised Mary. Another is the question mentioned in passing earlier (footnote 5): How are the actor’s volition and intention represented at the PES? As is well-known, causative events are usually, if not always, carried out by the actor volitionally and intentionally. They thus need to be represented in such a way as to differentiate prototypical (willful and deliberate) causative events from non-prototypical (inadvertent and unintentional) ones. The multi-tiered structure advocated above, therefore, is yet to be elaborated further. Still another question that awaits an answer concerns the issue touched upon
in footnote 3. The proper description of sentences with a certain type of adverb (e.g. *almost*) requires an event structure that properly identifies and represents a certain aspectual phase *before* a perfective point, but the PES put forward in this paper is not articulated in such a way. This poses the question of how it should be done. Yet another question to be addressed is one about the crosslinguistic applicability of the PES. If it is the proper representation of events, then the PES analyses of English verbal/adverbial expressions presented in this paper must be extended successfully to the corresponding expressions in other languages, too. It is worthwhile to see if and to what extent the PES can accommodate them. Given the results for Japanese (with which it fares quite well; recall (33) and see Miyakoshi (2008, 2009a, 2009b) for detailed discussion), it seems promising.

REFERENCES


Kageyama, Taro (2001) “Jo: Doosi Kenkyuu no Genzai (Introduction: The State of the Art of the Verb Study),” *Doosi no Imi to Koobun* (Verb Semantics and Con-


[received April 1 2009, revised and accepted November 28 2009]