ACADEMIC PRESS

Cognitive Psychology 45 (2002) 1–44

Cognitive Psychology

www.academicpress.com

Event templates in the lexical representations of verbs☆

Gail McKoon* and Talke Macfarland

Department of Psychology, Northwestern University, Evanston, IL 60208, USA

Accepted 14 August 2001

Abstract

Four experiments support the hypothesis that syntactically relevant information about verbs is encoded in the lexicon in semantic event templates. A verb's event template represents the participants in an event described by the verb and the relations among the participants. The experiments show that lexical decision times are longer for verbs with more complex templates than verbs with less complex templates and that, for both transitive and intransitive sentences, sentences containing verbs with more complex templates take longer to process. In contrast, sentence processing times did not depend on the probabilities with which the verbs appear in transitive versus intransitive constructions in a large corpus of naturally produced sentences.

1. Introduction

Shudder and *shake* describe similar kinds of movements, trembling vibrating sorts of movements, but they cannot be used in the same ways. While an entity can be said to shake and it can be said to shudder, and one entity can be said to shake another, one entity cannot be said to shudder

☆ Preparation of this article was supported by NIMH Grants R37-MH44640 and K05-MH01891 and by NIDCD Grant R01-DC01240. We thank Roger Ratcliff for contributions at all stages of this research.

* Corresponding author.

E-mail address: g-mckoon@nwu.edu (G. McKoon).

0010-0285/02/$ - see front matter
PII: S0010-0285(02)00004-X

another. *The trees shook*, *The trees shuddered*, and *The wind shook the trees* are all acceptable sentences, but *The wind shuddered the trees* is not (Atkins & Levin, 1995; Song, 1996). The question addressed in this article is what gives rise to constraints like these on verb usage: are the constraints syntactic or semantic?

One possibility is that the constraints are syntactic. Lexical information about *shake* would include the information that it can participate in both transitive and intransitive structures, whereas lexical information about *shudder* would include the information that it can participate only in intransitive structures. We propose and offer evidence for a different view, that the structures in which verbs can appear are determined by the verbs' meanings. It is because of a difference in their meanings that *shake* but not *shudder* occurs in transitive structures.

In the first sections of this article, we outline our proposal, which we label "meaning through syntax" (MTS). For verbs, "meaning through syntax" signifies that some of the meaning conveyed by verbs is expressed through and reflected in the syntactic structures in which the verbs appear. Insights into verbal meaning can be gained by examining syntactic structures. As candidate representations of the parts of verbal meaning relevant to syntactic structure, we propose "event templates," which we describe in detail below. The event template for a verb is a theoretical construct that is intended to explain in which syntactic structures the verb can and cannot occur. Furthermore, the parts of verbal meaning underlying syntactic behavior are also assumed to underlie sentence comprehension, and so event templates are expected to have observable consequences for comprehension. According to the MTS view, the representations of verbal meaning expressed in event templates should unite an explanation of a verb's syntactic behavior with its meaning and its role in sentence comprehension.

Below, we explain our reasons for adopting event templates as representations of syntactically relevant meaning. We then focus on an investigation of two specific classes of verbs, classes that provide particularly subtle tests of the MTS view. We show that both sentence production data and sentence comprehension data provide evidence for the psychological reality of the proposed event templates for the verbs of the two classes. The event templates successfully account for the sentence structures in which speakers do and do not produce the verbs, and they also allow accurate prediction of the relative degree of difficulty of comprehension for verbs of the two classes.

2. MTS event templates for verbs

When psycholinguists make reference to the lexicon, the lexical entry of a verb is understood to contain information about its orthography, its pro-

nunciation, and its meaning and also information about how the verb behaves in sentences. This latter information has frequently been divided into two parts: semantic information about how many and what kinds of arguments the verb relates (for example, an action verb like *shake* relates an agent argument and a patient argument) and syntactic information about the types of complements the verb can take (for example, verbs that can take noun phrase complements can be used transitively). This two-tiered approach to how verbs behave in sentences has often been suggested in psycholinguistics (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994, pp. 683–684, & citations therein). However, for MTS, we have rejected this approach. For MTS, information about how a verb behaves in sentences is expressed in a single, purely semantic, configuration—an event template—that shows the entities involved in the event denoted by a verb and the entities' relations to each other. Table 1 shows examples of four types of event templates, discussed in more detail below. The templates capture the MTS view that the sentence constructions in which verbs appear are semantically controlled and interpreted. It is these templates that we explore empirically in the latter parts of this article. Some of the reasons we have adopted event templates come from recent work in lexical semantics, which we review in the next section.

2.1. Background: Thematic argument roles and subcategorization frames

In most linguistic theories up to the early 1980s, verbal lexical entries contained three kinds of information: a verb's subcategorization frames, its thematic argument roles, and all of the rest of its meaning (e.g., Jackendoff, 1972). The subcategorization frames and argument roles were intended to jointly carry the parts of the verb's lexical entry relevant to syntax. For example, the argument role list for the verb *put* might have been something like *Agent Theme Location*, where the agent is the entity causing the event of putting, the theme is the entity being put, and the location is where it is put. The subcategorization frame would be (NP1) ((V) (NP2) (P NP3)),

Table 1
Event templates

Event template name	Template	Examples
State	*x* (IN STATE)	*x* (LOVE), *x* (EXIST)
Activity	*x* (ACT)	*x* (RACE), *x* (HIT)
Change of state	*x* (BECOME IN STATE)	*x* (BECOME ARRIVED), *x* (BECOME BLOOMED)
Externally caused change of state	α CAUSE (*x* (BECOME IN STATE))	α CAUSE (*x* (BECOME BROKEN)) α CAUSE (*x* (BECOME FADED))

where NP1 would take the subject position in a sentence, V is *put*, NP2 would take the direct object position, and P would head a prepositional phrase with NP3 in prepositional object position. For most verbs, there would have been multiple possible argument role lists and subcategorization frames. *Run*, for example, would have (at least) two argument role lists, *Agent* and *Agent Goal*, and the subcategorization frames (NP1) (V) and (NP1) ((V) (P NP2)), for sentences like *John ran* and *John ran to the store*.

The arguments of a verb were specified in both its argument role lists and its subcategorization frames. In linguistics, preference is given to theories that are general and explanatory, but at the same time as simple as possible. Redundancies, such as specifying a verb's arguments in both a role list and a subcategorization frame, are eschewed. General principles that constrain the form of grammatical representations, such as the Projection Principle (Chomsky, 1981, 1986), which ensures that lexical information is preserved in sentence structure, paved the way for the elimination of subcategorization frames. Moreover, researchers in linguistics came to believe that sentence structure "is largely predictable from the semantics of predicates" (Wasow, 1985, p. 202; also Bates & Goodman, 1997; Chomsky, 1986; Levin & Rappaport Hovav, 1996; Moens & Steedman, 1988; Pustejovsky, 1991; Tenny, 1994; Van Valin & LaPolla, 1997). Thus, in most linguistic theories, subcategorization frames were eliminated and thematic role lists alone mediated the mapping of arguments to positions in the syntax of a sentence. An important question is whether sentence structure is entirely predictable from semantic information. The research strategy employed by a number of linguists (e.g., Rappaport Hovav & Levin, 1998; Tenny, 1994) is to assume that it is, pushing meaning to predict as much of sentence structure as it possibly can.

Thematic role lists were purely semantic entities. They represented the parts of verbs' meanings that determined the syntactic sentence structures in which the verbs could appear. For this reason, they are referred to as expressing "syntactically relevant meaning." All the rest of a verb's meaning has been labeled "content" meaning (Grimshaw, 1993; Levin & Rappaport Hovav, 1995; Rappaport Hovav & Levin, 1998). We follow the convention of using these labels in this article.

With the elimination of subcategorization frames, problems with thematic role lists became apparent and they have been extensively discussed in the linguistics literature. Some of the problems are that there is no consensus as to the appropriate set of possible thematic roles or what their definitions should be, and there are no reliable criteria for determining which thematic role a given argument should bear (Rappaport Hovav & Levin, 1998). For example, should there be just one role to describe the arguments that appear as the objects of *study*, *eat*, and *see*, or should there be three different roles to describe these arguments (Croft, 1991, 1998; Dowty, 1989, 1991; Levin & Rappaport Hovav, 1996; Wechsler, 1995)?

Another problem is that thematic role lists cannot properly capture the relations among the different constructions in which a verb can participate and their concomitant variations in meaning. To illustrate this point, consider an example discussed by Rappaport and Levin (1988), the locative alternation, illustrated in the following sentences:

Bill loaded cartons onto the truck.
Bill loaded the truck with cartons.

Although positing a single list of thematic roles for the verb *load* in the two sentences (for example, *Agent Theme Goal*) would capture the near paraphrase relation between the sentences, it could not predict the variation in linking of the arguments to sentence positions: in the first sentence, the Theme *cartons* is realized as direct object and the Goal *truck* as prepositional object, and in the second sentence, the linking is reversed. Neither could a single role list explain the fact that we understand the truck to be full when it appears as direct object but not when it appears as prepositional object (Anderson, 1971; McKoon, Ratcliff, Ward, & Sproat, 1993). Of course, it could be assumed that there are two separate role lists for *load*, but this would lead to a proliferation of role lists, returning to the redundancy that the elimination of subcategorization frames was intended to avoid (Rappaport & Levin, 1988).

A further problem with role lists is that no consensus could be developed for how to map arguments to syntactic positions. Mapping was supposed to be accomplished through a thematic role hierarchy. A hierarchy listed roles in order such that, in many approaches, the highest ranking role was assigned to subject position in a sentence, the next highest to direct object position, and so on. But there was variation across theorists in both the choices of roles included in the hierarchy and in their relative rankings (see discussion in Croft, 1998; Levin & Rappaport Hovav, 1996). For instance, Grimshaw (1990) ranked Agent above Goal and Location, and Goal and Location equal, both above Theme. In contrast, Speas (1990) ranked Agent above Theme and Theme above the equally ranked Goal and Location, and Kiparsky (1985) ranked Agent above Goal, Goal above Theme, and Theme above Location. Overall, there was neither a universally accepted hierarchy nor a universally accepted set of rules for using a hierarchy to mediate between semantic arguments and syntactic structure. It became clear that, as Croft (1998, p. 29) pointed out, the problem was the lack of an underlying motivation for the ranking of the thematic roles because a hierarchy of isolated roles "does not take account of their relational nature."

Early versions of thematic role lists had no internal structure and a verb's arguments were distinguished only by their thematic role labels (e.g., Agent and Theme). Later representations introduced structure by distinguishing the external argument of a verb, which receives its thematic role assignment from the whole verb phrase (Williams, 1981), from internal arguments,

which are internal to the verb phrase. Among internal arguments, an argument that receives its thematic role directly from the verb was designated the direct internal argument and an argument that receives its thematic role from a preposition was designated an indirect internal argument (Marantz, 1984). External arguments were mapped to subject position in a sentence, direct internal arguments to direct object position, and indirect internal arguments to prepositional object positions. For the verb *put*, for example, Rappaport and Levin (1988) gave its representation as $x\langle y, \text{Ploc}\, z\rangle$, where x is the external argument, y the direct internal argument, and z the indirect internal argument. In this formulation, the syntactically relevant aspects of the representation were no longer the semantic content of the thematic roles, not Agent or Theme, but rather the representation's structure. Crucially, thematic role lists had been transformed into predicate decompositions.

2.2. Background: Predicate decompositions

Current research in lexical semantics retains from earlier work the separation of the content part of a verb's meaning from the syntactically relevant part that is expressed in a predicate decomposition and controls how the verb behaves in sentences. Different forms of predicate decompositions have been offered by, among others, Pustejovsky (1991), Rappaport Hovav and Levin (1998), Tenny (1994), and Van Valin and LaPolla (1997) (see also Hale & Keyser, 1993; Stevenson & Merlo, 1997; for a different type of structural representation). MTS event templates are most like Rappaport Hovav and Levin's (1998) decompositions.

It is usually assumed that there are four classes of verbs, each class with a different predicate decomposition. Table 1 shows the four classes in event template representations. The first template, for states, has a simple structure: x exists in some state. The template shows the entity, x (the "argument"), that participates in the event denoted by the verb and, with the semantic primitive IN STATE, it shows the state in which x is construed to be existing. This template represents the syntactically relevant parts of their meanings for verbs like *love* and *exist*. The second template, for activity verbs like *race* and *hit*, shows an entity, x, engaging in some activity, ACT. The third template shows the structure of simple change of state verbs like *bloom* and *arrive*. The semantic primitive BECOME specifies a change for the entity x to a new state. The fourth template is for verbs that we, following Levin and Rappaport Hovav (1995), call "externally caused change of state" verbs, verbs like *break*. The template shows two subevents, a causing event α and a change of state event, connected by the primitive CAUSE. The change of state subevent is a simple change of state x (BECOME IN STATE). The causing event α can be any of the four event types, a state, an activity, a simple change of state, or an externally caused change of state (Van Valin & LaPolla, 1997). It is typically represented in a sentence by an

entity engaging in the subevent; for example, *John* represents α in *John broke a window* and *the ball* represents α in *The ball broke a window.*

The four predicate decompositions shown in Table 1 are taken to define four types of events that verbs may denote. For any particular verb, its template is an instantiation of the verb into the appropriate template type, as is shown with the examples in the table, for instance, x (HIT) and x (BECOME ARRIVED). For all the verbs that share the same template, the assumption is that the syntactically relevant part of their meanings is the same. It is in the content parts of their meanings that they differ.

In much current work in lexical semantics, the four types of events are understood to correspond to the aspectual classes of verbs proposed by Vendler (1957) and Dowty (1979). Aspectual classifications sort verbs according to their inherent temporal properties. The three properties are telic/atelic, punctual/durative, and stative/dynamic. A telic event has an end point in time and an atelic event does not. A punctual event happens in an instant, whereas a durative event occurs over time. In a stative event, there is no change, whereas a dynamic event involves a change of some kind. The three properties sort verbs into four classes: States are said to be stative, durative, and atelic; activities are said to be dynamic, durative, and atelic; simple changes of state (often known as achievements) are said to be dynamic, punctual, and telic; and externally caused changes of state (often known as accomplishments) are said to be dynamic, durative, and telic.

The Vendler–Dowty fourfold distinction is attractive because there appear to be explicit diagnostics for evaluating verbs' behaviors with respect to the temporal properties that establish class membership. For example, manner of motion activity verbs such as *run* are said to denote atelic durative activities, while inherently directed motion verbs such as *arrive* are said to denote telic punctual achievements. Thus, *run* verbs, but not *arrive* verbs, can occur with the durative adverbial "for an hour" as follows:

The student ran for an hour.
*The student *arrived for an hour.*

Another example is the imperfective paradox, which is considered to distinguish between atelic verbs such as *run*, where the progressive entails the perfect, and telic verbs such as *arrive*, where such an entailment does not go through, as follows:

The student is running entails *The student has run.*
The student is arriving does not entail *The student has arrived.*

Despite its attractions and its widespread acceptance in lexical semantics, for the MTS view we do not equate event templates with aspectual classes. For one reason, it turns out that the classes of verbs defined by event templates are not the same as the classes defined by aspectual diagnostics. For example, *arrive* and *ascend* are both simple change of state verbs: for both

verbs, an entity changes state in that it moves from one location to another. But *arrive* is telic while *ascend* is atelic. Also, verbs can switch aspectual classes according to the context in which they are used. Hay, Kennedy, and Levin (1999) provided examples of verbs that show consistent switching patterns, verbs like *lengthen* which can switch from telic (*The tailor lengthened my pants*) to atelic (*The traffic lengthened my commute*).

Another reason for rejecting a correspondence between the classes of verbs defined by aspectual diagnostics and those defined by event templates is that the meaning embodied in event templates is not temporal in the sense defined for aspectual classes. Event templates describe events and the relations among them and their arguments. The relations show the placement of arguments in events and subevents, whether and how arguments are affected in the course of an event, and how subevents are connected to each other by causality. States and activities place arguments in events, simple changes of state place arguments in events in which the arguments undergo a change, and externally caused changes of state place arguments in two causally related subevents. The relations do not show the aspectual temporal nature of events. A change of state, for example, is a change of state whether it happens punctually or duratively. Thus, in the MTS view, event templates structurally define the relations of events and their arguments, not the aspectual temporal characteristics of events.

2.3. MTS event templates

It is important to understand what event templates are and what they are not. First, event templates are theoretical constructs designed to represent the portion of a verb's meaning that is relevant to sentence structure. A template does not, by any means, capture all the other parts of a verb's meaning. It is merely the skeleton of meaning upon which a sentence can be built. A verb, through its event template, can be said to structure the meaning of a sentence (Tenny, 1994).

The meaning in a verb's template is a reflection of the other parts of its meaning. *Run* has an activity template because it is an activity; *break* has an externally caused change of state template because someone or something causes something else to break. But the meaning in a verb's template is not *exactly* predictable from the other parts of the verb's meaning. Borrowing an example from Levin and Rappaport Hovav (1995), the verb *blush* is construed differently across languages. It is construed as an activity (x (ACT)) in English and Dutch, but it is construed as a simple change of state (x (BECOME IN STATE)) in Italian. In English and Dutch, it can be said that someone blushed *for 10 minutes*, indicating an activity. But we do not say that someone blushed *in a minute*, meaning it took a minute for the person to change from the nonblushing to the blushing state. In Italian, the use of the verb is exactly reversed: blushing for 10 minutes is an unacceptable

use of the verb and blushing in a minute is acceptable (see Levin & Rappaport Hovav, 1995, for further discussion).

The *blush* example illustrates a point about event templates that cannot be overstressed: Event templates are *not* intended to represent events in the world. Instead they express construals of events in the world; that is, they express how speakers use language to conceptualize events (Grimshaw, 1993; Levin & Rappaport Hovav, 1995; Smith, 1991). Apparently, for Italian speakers, the verb *blush* structures an event of blushing as a change of state, while for English and Dutch speakers, *blush* structures the event as an activity.

Unlike thematic role lists, the entities represented in event templates (e.g., the *x* in *x* (ACT)) are simply place holders. There is nothing semantic about them in themselves and they are not given meaning by labels such as "agent" or "patient." The meaning of a template lies in the way it structures entities and the events in which they participate.

The entities that are represented in an event template are only those that are required for expression in a sentence. The verb *erode*, for example, has the template *x* (BECOME ERODED), and the entity that changes state (*x*) appears in any sentence that uses the verb *erode*. *Erode* is also often used in a transitive sentence (*High tides eroded the beach*) but a second argument (*High tides*) is not required and is not represented in the template. It is assumed that optional arguments like *High tides* are optionally added to a sentence from the content parts of a verb's meaning.

This is a point of sharp contrast between event templates and the thematic role lists used in some current psycholinguistic research. Typically, in a role list approach, a verb's lexical entry would have encoded with it all the different possible role lists with which the verb could occur (e.g., for *erode*, one role list for sentences like *The beach eroded* and a different role list for sentences like *High tides eroded the beach*). In the MTS approach, the lexical entry for a verb has only a single event template (unless the verb is truly ambiguous), to which optional arguments can be added.

Event templates are compositional in that more complex templates can be composed of simpler ones. For example, the *x* (ACT) template for *sweep*, which would underlie the sentence *John swept*, can be combined with a change of state template, *x* (BECOME IN STATE), and the primitive CAUSE to produce an externally caused change of state template for the sentence *John swept the leaves into a pile*. However, it is usually hypothesized in lexical semantics that there is a limit on the complexity of event templates such that the most complex template possible is the externally caused change of state template (Rappaport Hovav & Levin, 1998).

For thematic role lists, the semantics–syntax interface is accomplished via a fixed hierarchy of roles, as discussed above. For event templates, projection from event templates to sentence structures is implemented by linking rules. Linking rules can be illustrated with two rules suggested by Levin

and Rappaport Hovav (1995). The "Immediate Cause" linking rule states that the argument referring to the immediate cause of an event denoted by a verb will take the subject position in a sentence. For example, the immediate cause for *John broke a window* is something John did, α, and so α is placed in subject position, where it is referenced by *John*. The "Directed Change" linking rule states that the argument of a template that corresponds to the entity undergoing a directed change takes direct object position. In *John broke the window*, it is the window that undergoes directed change and so it is the direct object. We discuss linking rules in more detail later, but what is important to note here is that the linking rules are formulated in semantic terms; they rely on an understanding of the eventuality denoted by a verb, the participants therein, and their relations to one another, as laid out in the verb's event template.

3. Internally caused and externally caused change of state verbs

MTS is a proposal that is new to psycholinguistics. Its strongest value is that it generates hypotheses about sentence production and comprehension that would not arise from other current frameworks. Because it is so new, we cannot yet provide a large amount of supporting evidence and so it is essential to choose empirical tests that are as stringent as possible. The two classes of verbs studied in depth in this article were chosen because they share elements of meaning and they appear in many of the same sentence structures, yet they have different event templates.

The two classes of verbs are internally caused change of state verbs, known by their prototypical member as "bloom" type verbs, and externally caused change of state verbs, known as "break" type verbs (Levin, 1993; Levin & Rappaport Hovav, 1995; Smith, 1970). Internally caused change of state verbs are a subclass of simple change of state verbs, and so they have the template x (BECOME IN STATE). Intuitively, an internally caused change of state is a change of state for which the cause resides in the entity undergoing the change. For example, flowers bloom because of something internal to flowers. An externally caused change of state, on the other hand, comes about because of something external to the entity that undergoes the change of state. The template α CAUSE (x (BECOME IN STATE)) shows an external causing event and a change of state. For example, an entity does not break solely because of its own properties, even though it is true that it must have certain properties in order for it to be breakable (Levin & Rappaport Hovav, 1995). The responsibility for windows breaking lies not with windows themselves but with some external force.

Verbs of these two classes sometimes appear similar in meaning (which is why we thought them a particularly stringent test for MTS). For example, *shrivel* is classified as an externally caused change of state, whereas *wither*

is classified as an internally caused change of state (Levin & Rappaport Hovav, 1995). This raises the question of how the decision is made whether to classify a verb as externally or internally caused. Why should *shrivel* have the α CAUSE (x (BECOME IN STATE)) template while *wither* has the x (BECOME IN STATE) template?

Originally, the distinction between internal and external causality was motivated by linguistic intuition (Levin & Rappaport Hovav, 1995; Smith, 1970) and by the fact that it allowed a good fit between the verb classes and larger theoretical issues in lexical semantics (e.g., Levin & Rappaport Hovav, 1995), which are not of relevance here. Levin and Rappaport Hovav (1995) proposed two diagnostics to buttress the distinction. The first was that externally but not internally caused change of state verbs can appear in transitive sentences. Consistent with this diagnostic, a person can be said to break a window (external causality) but he cannot be said to bloom a rose (internal causality). Levin and Rappaport Hovav (1995) pointed to the event templates for the two classes of verbs as an explanation of the diagnostic: the internally caused changed of state verbs cannot appear in transitive sentences because their template, x (BECOME IN STATE), has only the one argument x; in contrast, the externally caused change of state verbs can appear in transitive sentences because their template, α CAUSE (x (BECOME IN STATE)), has two subevents, one for each of the arguments in a transitive sentence.

The second diagnostic suggested by Levin and Rappaport Hovav (1995) was that the range of entities for which a change of state can be denoted by an internally caused change of state verb is smaller than the range for which a change of state can be denoted by an externally caused change of state verb. The intuition was that only flowers and other plants can be said to bloom, whereas all manner of things can be said to break. Again, Levin and Rappaport Hovav (1995) explained this with appeal to the verbs' event templates. Because responsibility for an internally caused change of state lies within the entity changing state, only those entities with the requisite internal characteristics can undergo the change of state process denoted by an internal causation verb.

These two diagnostics express precisely the research plan we outlined at the beginning of this article. The goal is to understand the constructions in which verbs appear in semantic terms. If internally caused verbs do not appear in transitive sentences and if they are used to denote a more limited range of entities changing state than externally caused verbs, then we look to their meanings—to their event templates—for an explanation because their behaviors in sentences are a reflection of those meanings.

We began our research with the aim of testing whether the two diagnostics just described accurately distinguish internally from externally caused change of state verbs. It turned out that they do not (McKoon & Macfarland, 2000). McKoon and Macfarland studied large numbers of naturally

produced sentences in which the verbs occurred. The picture that emerged was more complex than the two diagnostics would suggest, and the reflections of syntactically relevant meaning in the structures of the sentences in which the verbs of the two classes were used were much more subtle. Nevertheless, the results of the studies gave strong support for the MTS event templates proposed for the two classes of verbs. In the next sections of this article, we describe McKoon and Macfarland's data and show how the event templates account for the data. We begin by explaining how arguments in the templates of the two verb classes are linked to positions in sentences.

3.1. Linking rules for internally and externally caused change of state verbs

First, consider internally caused change of state verbs in intransitive sentences such as *The beach eroded*. When a verb of this class has only a single argument, call it *x*, *x* is the entity that both undergoes the change of state and is responsible for the change of state. When beaches erode, for example, they change state in that they become eroded and they are also responsible for the change in that their physical makeup allows erosion. This means that the Immediate Cause and Directed Change linking rules both apply, but they are in conflict: By the Immediate Cause rule, the entity responsible for the change of state, which is *x*, should take subject position but by the Directed Change rule, the entity affected by the change, which is also *x*, should take direct object position. However, the Directed Change linking rule is assumed to take precedence over the Immediate Cause linking rule (Levin & Rappaport Hovav, 1995), so *x* is assigned the direct object position. As *x* is the only participant in the event, it moves to subject position (a move well-established in the linguistics literature; e.g., Burzio, 1986; Perlmutter, 1978). Thus, in *The beach eroded*, *beach* is the entity undergoing erosion, and in the sentence, it is an object moved to subject position.

Now, consider what happens when there is a second participant in the change of state event, call it *y*. Although an internally caused change of state verb has only one argument in its template, *x*, another argument can be added from the content part of the verb's meaning, and so the verb can appear with two arguments. We can say, for example, *High tides eroded the beach*, where *x* is the beach and *y* is high tides. However, because there is only one event in *erode*'s template, *y* must be a participant in that event. This has the important consequence that *y* should be an entity that is intrinsically involved in the internally caused change of state process. High tides and strong winds and ocean currents are all forces that can be involved in the process of beach erosion. Other kinds of entities cannot be; we do not say *John eroded the beach* or *The shovel eroded the beach*. When there are two arguments for an internally caused verb, *x* in the template and *y* from the content part of the verb's meaning, they are linked to direct object and

subject positions, respectively. The affected entity, *x*, does not have to move to subject position as it would if it were the only argument so it remains in object position, leaving subject position for *y* (McKoon & Macfarland, 2000).

To review, internally caused change of state verbs can appear in both intransitive and transitive constructions. In intransitive sentences, the participant in the event that is both responsible for the change of state and affected by it appears in subject position (moved from direct object position). In transitive sentences, the entity responsible for the change of state and affected by it remains in object position; the second participant is in subject position; and because there is only the single change of state event in the template, the second participant is intrinsically involved in that change of state. When Levin and Rappaport Hovav (1995) suggested that internally caused change of state verbs could appear only in intransitive sentences, what they overlooked was the possibility of a second argument being added from the content part of the verbs' meanings (something they later suggested for other classes of verbs; Rappaport Hovav & Levin, 1998). As we show below, for some internally caused change of state verbs, the probability with which they appear in transitive sentences is quite high, for some verbs as high as 0.60.

When externally caused change of state verbs appear in transitive constructions, the entity undergoing the change of state is in object position (by the Directed Change linking rule). The second participant is licensed by the causing subevent of the externally caused change of state, and it appears in subject position (by the Immediate Cause linking rule). In the template, the causing subevent is separate from the change of state subevent and therefore not constrained by it. Externally caused changes of state can be conceptualized as having a wide variety of causes. A window can be broken by a person, by a shovel, by high tides, and many other things. When externally caused change of state verbs appear in intransitive constructions, the entity undergoing the change of state is in object position, moved to subject position (just as for intransitive internally caused change of state verbs).

In sum, both internally caused and externally caused change of state verbs can appear in both transitive and intransitive structures. The linking rules map their different templates onto the sentence structures. We stress that the linking rules, as proposed, are abstract linguistic descriptions. It is an open question whether or how they might be reflected in psychological processing.

The templates combined with the linking rules led McKoon and Macfarland (2000) to propose a new diagnostic for distinguishing the internal from the external causation verbs. Because the template for the internal causation verbs contains only one event, the subject of a transitive sentence should be an inherent participant in that event. This constraint does not apply to the external causation verbs because their templates contain two separate sub-

events. Following this reasoning, McKoon and Macfarland (2000) proposed that the kinds of entities that appear in the subject positions of transitive sentences should be more constrained for the internally caused than the externally caused change of state verbs.

3.2. Distinguishing internally and externally caused change of state verbs

To test the diagnostic just described, as well as the two diagnostics proposed earlier by Levin and Rappaport Hovav (1995), McKoon and Macfarland examined language in natural use. Their studies relied on a corpus of about 180 million words of naturally produced English text. The corpus is made up of text from newspaper and magazine articles, adult and juvenile fiction and nonfiction, and spoken dialogue (e.g., "Larry King Live"). The verbs in McKoon and Macfarland's studies included all of the internally caused change of state verbs listed in Levin and Rappaport Hovav (1995) for which there were sufficiently many corpus tokens, except for the verb *burn*, which we believe to be ambiguous in that it has both internally caused and externally caused change of state meanings. The verbs in McKoon and Macfarland's studies are the 14 verbs listed in Table 2 as those used in Experiment 2.

For each of these verbs, McKoon and Macfarland (2000) examined an average of 200 sentences from the corpus. For some of the internally caused verbs, those in the left-hand column, the mean probability with which they occurred in transitive sentences was only .06, but for the verbs in the right-hand column, the mean probability was .45. All 14 of the verbs occurred at least twice in transitive constructions, even *bloom* (e.g., *. . .it* (*a shrub*) *blooms white flowers in summer*). For *corrode*, *erode*, and *ferment*, the probability was especially high, .63, .67, and .54, respectively. Salt water, sweat, and acid rain all are said to corrode things; rivers, rain, and rising tides all are said to erode things; and microorganisms are said to ferment things. It is these data that led McKoon and Macfarland (2000) to reject Levin and Rappaport Hovav's (1995) proposal that internally caused change of state verbs can be distinguished from externally caused change of state verbs because only the latter appear in transitive sentences.

Examination of naturally produced corpus sentences also led to rejection of Levin and Rappaport Hovav's (1995) second diagnostic for distinguishing internally from externally caused change of state verbs. They had suggested that, for internally caused change of state verbs, the range of entities denoted as changing state should be limited. The entities changing state are the subjects of intransitive sentences and the objects of transitive sentences. This diagnostic fails because of the power of speakers to use verbs to place their own construals on events. Not only are flowers said to bloom, but also peace, beauty, and football players. Beaches are said to erode but it can also be said that the stock market, good will, and authority erode. For

Table 2
Externally caused and internally caused change of state verbs used in experiments 1, 2, 3, and 4

Externally caused verbs α CAUSE (*x* (BECOME IN STATE))				Internally caused verbs *x* (BECOME IN STATE)			
Low probability transitive	Expt.	Higher probability transitive	Expt.	Low probability transitive	Expt.	Higher probability transitive	Expt.
abate .10	1–4	dissipate .41	1–4	deteriorate .01	1–3	blister .22	1–3
atrophy .03	1–3	fossilize .60	1 and 2	germinate .06	1–4	corrode .63	1–4
crumble .05	1–3	fray .52	1–4	rot .08	1–3	erode .67	1–4
explode .07	1–3	redden .24	1–4	stagnate .02	1–4	ferment .54	1–4
shrivel .11	1, 2, and 4	thaw .61	1–4	wilt .06	1–4	sprout .26	1–4
vibrate .03	1–4	oxidize .21	1 and 3	wither .12	1–4	swell .37	1–4
awake .05	2 and 3	splinter .49	2	bloom .00	2 and 3		
fade .01	2 and 4	stiffen .62	3 and 4	rust .14	2 and 3		
proliferate .02	3	scorch .85	4	blush .00	4		
mellow .15	3 and 4						

Note. With each verb is given the probablity with which it occurs in a transitive construction in the corpus of naturally produced sentences.

sentences from the corpus, McKoon and Macfarland (2000) examined the subjects of intransitive sentences and the objects of transitive sentences for the 14 internally caused verbs mentioned above and for 14 externally caused change of state verbs (listed by Levin & Rappaport Hovav, 1995) matched on Kučera–Francis frequency (the verbs marked for Experiment 2 in Table 2). No significant differences were found in the ranges of entities described as changing state by internally versus externally caused change of state verbs.

According to the diagnostic proposed by McKoon and Macfarland (2000), internally and externally caused change of state verbs can be distinguished by the subjects of the transitive sentences in which they appear. The diagnostic is that the entities appearing in subject position in transitive sentences should be more constrained for the internally than the externally caused verbs. The reason is that there is only the one, change of state, event in the template of an internally caused change of state verb and so a second participant in that event (the participant in subject position in a transitive sentence) must be an inherent participant in the change of state event. It is not possible to provide an exact rule for determining whether an entity is an inherent participant in a change of state event but, if the entities that change state are limited to concrete entities, then there can be good agreement among observers. Internally caused change of state verbs mostly describe natural processes (see the verbs in Table 2), so when the entities that undergo internally caused changes of state are concrete entities, the optional participants are mostly natural forces. In *High tides eroded the beach*, erosion is a natural process and high tides are a natural force. Similarly, for flood waters rotting carpet, heat wilting flowers, and sun blistering paint, there is good agreement that flood waters, heat, and sun all represent natural forces that participate in the rotting, wilting, and blistering changes of state. McKoon and Macfarland's (2000) corpus data showed that, in accordance with prediction, for corpus sentences with concrete objects, 80% of their subjects were natural entities like those just mentioned, entities that can be construed as inherent participants in the denoted events. The other 20% were split across three other categories: human-made artifacts, animate beings (people and, rarely, animals), and abstract entities. These latter cases appear to be either extensions of causality to humans (people are often said to ferment alcoholic beverages, even though microorganisms are actually doing the fermentation) or metaphorical uses of the verbs (e.g., *jealousy corroding a face*).

In contrast, and again in accordance with prediction, for corpus sentences with externally caused change of state verbs, the subjects were not predominately of any one category but instead were divided across all four categories: natural forces, human-made artifacts, animate beings, and abstract entities. Thus, the naturally produced sentences from the corpus showed a constraint (statistically significant) on the subjects of transitive sentences for internally but not externally caused change of state verbs.

When the entity undergoing an internally caused change of state in a transitive sentence is an abstract entity, McKoon and Macfarland (2000) found that almost anything can be conceptualized as a contributing factor, inherent to the change of state process. When abstract entities change state, speakers and writers can use verbs to place their own particular construals on events. For example, many kinds of things can be construed as being involved in *rotting the minds of the young*, ranging from natural substances to other people to behaviors. A corpus sentence said that movies, which we classify as artifacts, *rot the minds of the young*. In referring to the effect of movies on young minds as rotting, the writer uses an internal causation verb; that is, the writer conveys a metaphorical, internally caused, process that is abetted by the negative qualities of movies. Other examples of transitive sentences with internally caused change of state verbs with abstract objects are *sweltering heat wilted Republican enthusiasm* and *the Israelis can ferment their schemes against Damascus*. Thus, for abstract entities, no difference was expected between the transitive subjects of internally versus externally caused change of state verbs, and no statistically significant difference was found.

McKoon and Macfarland's conclusion was that an externally caused change of state verb can be distinguished from an internally caused change of state verb by examination of entities in subject position in transitive sentences with concrete objects. The classifications of verbs made on this basis matched the original classifications that were based on linguistic intuitions (Levin & Rappaport Hovav, 1995) for all those verbs for which there were sufficient data (i.e., a sufficient number of transitive sentences in the corpus).

3.3. Summary

According to the MTS view, the meaning encoded in the lexical entry for a verb can be divided into two parts, syntactically relevant meaning that determines how a verb structures its arguments in a sentence and the rest of the verb's meaning (its content meaning). Syntactically relevant meaning is represented by event templates that lay out the arguments of a verb and how the arguments participate in events. The event templates are semantic in nature. Linking rules project arguments to sentence positions, both the arguments in templates and optional arguments that come from the content parts of verbs' meanings.

The event template for a verb indicates how an event described by the verb is to be construed. An event described by the verb *bloom* is to be construed as an internally caused event. Of course, we could imagine how blooming could have external causes (e.g., a gardener's efforts), but the theoretical claim is that the event template does not incorporate all the variety of ways we could imagine describing an event denoted by the verb. Instead, the verb places a particular construal on the event—internal causality.

It is important to emphasize the point that verbs give construals of events. Any given event may be described in many ways, but a speaker or writer chooses a verb to present one particular conceptualization of the event. McKoon and Macfarland (2000) made this point with an example from the corpus about the sad state of the pyramids. This state might be construed as internally caused and so described as *deteriorating* or it might be construed as externally caused and so described as *crumbling*. Also, individual verbs, like *deteriorate* and *crumble*, can be used to describe events involving many different kinds of entities—people, society, morals, minds, and the quality of education can all be construed as *deteriorating* (internally caused) or *crumbling* (externally caused).

Given this flexibility in verb usage, it might be thought surprising that a diagnostic based on naturally produced sentences can distinguish internally from externally caused verbs. As just exemplified with *deteriorate* and *crumble*, the differences between verbs of the two classes are often subtle. Nevertheless, the corpus data show that speakers and writers do, statistically, obey event template constraints on verbs of the two classes. The difference between verbs of the two classes shows up not in the entities that are described as changing state, but in the entities that take subject position in transitive sentences. For internally caused changed of state verbs, the entity undergoing the change of state is also responsible for it. Thus, when an internally caused change of state verb is used in a transitive sentence, the entity taking subject position should be an inherent participant in the change of state event. That this is true for internally but not externally caused change of state verbs was demonstrated with transitive sentences with concrete objects.

4. Comprehension experiments

The central MTS hypothesis for the research described here is that verbs and the sentences in which they appear are understood through processes that make use of the verbs' underlying lexical semantic representations, that is, their event templates. For the two classes of verbs under investigation, the hypothesis is that comprehension of externally caused change of state verbs and the sentences in which they appear involves understanding that there is a change of state and that it comes about because of an external cause, whereas comprehension of internally caused change of state verbs and their sentences involves understanding only a change of state. Because the templates of externally caused change of state verbs are more complex than those of internally caused change of state verbs, we expect them to take more time to comprehend.

In the MTS view, there is only a single basic event template for each verb (unless it is truly ambiguous). Therefore, it is also hypothesized that the lexical semantic representation of a verb laid out in the event template is a part

of the verb's meaning that is used by comprehension processes whenever the verb is encountered. Whatever the syntactic construction in which a verb is used, for example, whether the verb is used transitively or intransitively, a sentence with an externally caused change of state verb is understood as involving both the causing and the change of state subevents, whereas a sentence with an internally caused change of state verb is understood as involving only a change of state.

McKoon and Macfarland (2000) and McKoon and Ratcliff (submitted for publication) provided initial support for the MTS hypotheses. The psychological reality of the claim that the lexical semantic event structures of external causation verbs are more complex than the lexical semantic event structures of internal causation verbs was tested with simple sentences constructed for verbs of the two classes. In one experiment, the sentences were intransitive (e.g., *The wind abated*) and in a second experiment, the sentences were transitive (e.g., *The severe drought withered the crops*). In both experiments, subjects were asked to judge, yes or no, whether the sentences were acceptable. Response times were significantly shorter for the sentences with internal causation verbs than the sentences with external causation verbs in both experiments.

In this article, we provide four new experiments that strengthen and considerably broaden support for the MTS claims. For the first experiment, the sentences were written to be unacceptable, such as *Some painters germinate medals*, and subjects were asked to judge the acceptability of the sentences. The MTS prediction was that judgments would be faster when the sentences contained internal causation verbs. The use of unacceptable sentences like *Some painters germinate medals* has the advantage that verbs of the two classes, internal causation and external causation, can be used in the same sentence frames. Compared to the use of acceptable sentences, which must be different for verbs of the two classes, this considerably minimizes the risk of confounding variables. In the second and third experiments, subjects were asked simply to read acceptable sentences, intransitive sentences in the second experiment and transitive sentences in the third experiment. These experiments extended McKoon and Macfarland's (2000) findings from the acceptability judgment paradigm to the more commonly used reading time paradigm. In the fourth experiment, verbs of the two classes were tested in lexical decision. If event templates are so important to verbs' lexical entries that differences in their complexity affect sentence comprehension time, then the differences might also be expected to show up in lexical decision response times. Furthermore, a difference in lexical decision response times would indicate that the internal–external effect was not dependent on a verb being in a sentence, that is, not dependent on interactions between the verb and other constituents of a sentence.

In all four experiments, the more complex templates for the external causation verbs were expected to lead to longer response times. This could

occur for any of several reasons. For example, the external causation template has two subevents and the internal causation template only one; the external causation template has the CAUSE relation and the internal causation template does not, and so on. We outline a variety of possible processing mechanisms under Section 9.

MTS differs sharply from some constraint-based models of sentence comprehension. For these models, the lexical representation of a verb lists all the different semantic or syntactic structures in which the verb can appear along with the probability of occurrence of each structure. The assumption is that sentences with structures that occur more frequently for the verb are processed faster than sentences with structures that occur less frequently. For MTS, there is only a single event template underlying all uses of a verb and lexical entries do not encode frequency of occurrence information. The sentences in Experiments 1, 2, and 3 were transitive and intransitive sentences, and so we thought it possible that the relevant information encoded in a constraint-based lexicon might be the probabilities with which the verbs are used in transitive versus intransitive constructions. Therefore, for both the internally and externally caused change of state verbs, we included in the experiments verbs with a high probability and verbs with a low probability of occurring in transitive sentences.

4.1. The verbs used in the experiments

The verbs used in the experiments are shown in Table 2. For each verb, we calculated the probability that it appeared in a transitive construction in sentences in the corpus. The total number of verbal tokens examined for each verb ranged from 33 to 587, with an average of 200; for only three verbs was the number examined less than 100. Each token was classified as transitive (including passives), intransitive, adjectival, or nominal. The transitive probability for each verb was calculated as the ratio of the number of transitive tokens divided by the number of transitive tokens plus the number of intransitive tokens.

Given the internally caused change of state verbs in Table 2 and their probabilities of being produced in transitive sentences, we chose externally caused change of state verbs to match them. The externally caused verbs were chosen from the verbs classified as externally caused by Levin and Rappaport Hovav (1995) and they were matched to the internally caused verbs in terms of Kučera–Francis frequency and in terms of transitive probability (calculating transitive probability from corpus sentences in the same manner as for the internally caused change of state verbs).

Table 2 shows the internally and externally caused verbs grouped by their transitive probabilities. Over both classes, the mean transitive probability for the low probability verbs is .06, and for the higher probability transitive verbs, the mean is .45. For both classes, the transitive and intransitive tokens

combined make up the large majority of verbal tokens from the corpus. For the internally caused verbs, only 8% of the tokens were adjectival and 5% nominal, and for the externally caused verbs, only 7% were adjectival and 3% nominal.

The internally and externally caused verbs in Table 2 were also matched on several other dimensions: The mean numbers of syllables per verb were 1.8 and 2.0, for the internally caused and externally caused verbs, respectively. We measured the verbs' imageability via ratings; subjects rated each verb on a scale of 1 to 5 with 1 being most imageable. The mean ratings were 2.3 and 2.5, respectively. All of the verbs except *blister* and *splinter* occur more frequently as verbs than nouns in the Francis and Kučera (1982) norms. We also checked the numbers of different senses listed for each verb in WordNet (Fellbaum, 1998). It might be thought that verbs in one or the other of the two classes have more variations in meaning, that is, more different senses. WordNet lists all the senses researchers in the WordNet project have found for each verb. The mean numbers of senses per verb were 2.1 and 2.6 for the externally caused and internally caused verbs, respectively. For all these dimensions, the differences between the two classes of verbs were not significant, F's < 1.4. There were also other dimensions, specific to each individual experiment, to be matched between the two sets of verbs. The sets of verbs used in each of the experiments were slightly different in order to achieve the best matches we could on these other dimensions.[1] For each verb, Table 2 shows in which experiments it was used.

5. Experiment 1

In Experiment 1, subjects were asked to judge sentences as "acceptable" or "unacceptable." There were unacceptable sentences for each of 12 externally caused change of state verbs and 12 internally caused change of state verbs. Six of each set of 12 verbs were low probability transitive and 6 were higher probability transitive. According to MTS, judgment times should be slower for sentences with externally caused change of state verbs than for sentences with internally caused change of state verbs, but judgment times should not be significantly affected by the probabilities that the verbs occur transitively versus intransitively in naturally produced sentences.

[1] Across the experiments, verbs and sentences were matched on a number of dimensions (imageability, plausibility, preexperimental semantic associations, and subjective frequency) that required ratings from subjects. For each rating study, the words or sentences of interest were mixed with filler items (about twice as many fillers as items of interest) in random order (one random order for half the subjects, and a different random order for the other half). In each rating study, there were from 10 to 12 subjects, all Northwestern University undergraduates participating in the study for course credit.

5.1. Method

Materials. The verbs for the sentences used in the experiment were those shown in Table 2 for which Experiment 1 is indicated. The mean Kučera-Francis frequency was 4.8 for the 12 external causation verbs and 5.8 for the 12 internal causation verbs. For all the verbs used in the experiments in this article, the reported frequency is the total frequency of verb uses given by Francis and Kučera (1982). For six of the external causation verbs, the mean probability of appearing in a transitive construction in the corpus is .07 and for the other six, it is .43. For the internal causation verbs, the corresponding means are .06 and .45. All of the experimental sentences had people as their subjects, so the externally and internally caused verbs were matched in the probabilities with which they appear in the corpus with people as their subjects; the mean probability was .12 for the externally caused verbs and .11 for the internally caused verbs.

Unacceptable sentences were constructed for the verbs, each sentence of the form *Some noun verb noun*. The subject noun was some kind of person (e.g., *hunters* and *tourists*) and always plural. The verb was used in its present tense. The object noun of the frame was always a thing, always plural. Six sets of four sentence frames per set were constructed. The nouns for the four sentence frames of a set were approximately equated for Kučera–Francis frequency. The four frames of a set were used with two external causation verbs and two internal causation verbs, with assignment of verbs to sentence frames counterbalanced across subjects. An example of one of the sets of frames with the verbs assigned to it is *Some painters abate/atrophy/germinate/stagnate medals*, *Some athletes abate/atrophy/germinate/stagnate brushes*, *Some hunters abate/atrophy/germinate/stagnate clocks*, *Some tourists abate/atrophy/germinate/stagnate needles*. The complete set of sentences is given in Appendix A.

Filler sentences were also constructed, 80 of them acceptable and 32 of them unacceptable, all four words in length.

Procedure. All stimuli were displayed on the screen of a PC monitor. The experiment began with 35 lexical decision test items, used for practice with the PC keyboard. Then the experimental and filler sentences were displayed in 12 lists of 9 sentences per list, preceded by 1 practice list. At the beginning of each list, subjects pressed the space bar to begin the list. Then the sentences were displayed one at a time. Subjects were to read each sentence and press the "?/" key if it was an acceptable sentence in their judgment, the "z" key if it was not. After a key press, the sentence was cleared from the screen and there was a 500 ms pause before the next sentence.

Each list of nine sentences was made up of one experimental sentence with an external causation verb and one with an internal causation verb, plus five acceptable and two unacceptable filler sentences. Each experimental sentence was presented only once to each subject (and this was true for

all the experiments described in this article). Each of the experimental sentences was immediately preceded by an acceptable filler sentence. Otherwise, the sentences of a list were presented in random order, with the random order changed after each second subject.

Subjects. There were 20 Northwestern University subjects, each participating in the experiment for credit in an Introductory Psychology course.

5.2. Results

Means were calculated for judgment times and response probabilities for each subject in each of the experimental conditions, and means of these means are shown in Table 3. For filler sentences, the mean judgment time for "yes, acceptable" responses to acceptable sentences was 1792 ms (.91 probability of a "yes" response) and the mean for "no, unacceptable" responses to unacceptable sentences was 2158 ms (.81 probability of a "no" response).

For sentences with external causation verbs, judgment times were significantly longer than for sentences with internal causation verbs, $F1(1,16) = 6.1$ $(MSE = 19{,}419)$ and $F2(1,10) = 6.7$ $(MSE = 15{,}964)$ ($p < .05$ throughout this article; $F1$ reports analyses with subjects as the random variable, and $F2$ analyses with items as the random variable). (The $F1$ ANOVA included counterbalancing as a factor; the interaction between verb type and this factor was not significant. The $F2$ ANOVA included the counterbalancing factor, and external and internal causation verbs that appeared in the same sentence frame were paired.) The standard error of the judgment time means was 31 ms.

Table 3 shows the data divided according to whether the verbs in the sentences have low (.06 on average) or higher (.44 on average) probabilities of appearing in transitive constructions in the corpus. First, these data show that the external–internal difference appeared at both levels of probability. For the two levels of probability, there was no significant difference in the size of the external–internal effect (F's < 1.0). Second, judgment times were shorter for sentences with higher probability of transitive usage than for sentences with lower probability of transitive usage, an effect significant with

Table 3
Data from Experiment 1

	Externally caused α CAUSE (x (BECOME IN STATE))		Internally caused x (BECOME IN STATE)	
	Judgment time (ms)	Pr. "no"	Judgment time (ms)	Pr. "no"
All sentences	1885	.92	1752	.92
Low probability transitive	1919	.96	1825	.95
Higher probability transitive	1852	.89	1679	.90

the subjects' analysis of variance, $F1(1,16) = 7.9$ ($MSE = 90{,}943$), but not the items analysis, $F2(1,10) = 2.0$ ($MSE = 15{,}964$). However, this effect is qualified by a speed–accuracy trade-off: not only did the sentences with the verbs with the higher probability of transitive usage have shorter judgment times, they also were associated with a lower probability of "no" responses, $F1(1,16) = 6.1$ ($MSE = .02$), $F2(1,10) = 7.5$ ($MSE = .004$).

We checked that the external–internal difference held for those of the verbs that have low Kučera–Francis frequencies. For the verbs with Kučera–Francis frequencies at the mean of the verbs we used or lower (i.e., less than 6), the external–internal difference was still large, 1851 ms (.93 probability of "no" response) versus 1797 ms (.93 probability of "no" response), respectively.

We also checked that the external–internal difference held for verbs that have a low probability of occurring with people subjects in the corpus. The experimental sentences always had people as subjects, and a possible hypothesis is that a sentence with people as a subject and a verb that rarely takes people as a subject might be so easy to reject that the external–internal difference disappears. However, this was not the case. For the half of the verbs of each class for which the probability of people as a subject was .10 or less (i.e., below the mean probability), judgment times for external causation verbs averaged 1944 ms (.95 probability of "no" response) and judgment times for internal causation verbs averaged 1781 ms (.93 probability of "no" response).

6. Experiment 2

In Experiment 1, sentences with internal causation verbs were processed faster than sentences with external causation verbs, as expected from the MTS view. This result confirms findings by McKoon and Macfarland (2000). In McKoon and Macfarland's (2000) experiments, the materials were acceptable intransitive sentences and acceptable transitive sentences, for example, *The flowers bloomed* and *The salt water rusted the machinery*. For both the intransitive and the transitive constructions, sentences with internal causation verbs were processed more quickly. McKoon and Macfarland matched the sentences for the two types of verbs on several dimensions: number of words, Kučera–Francis frequencies of the verbs and the other words of the sentences, and probability with which the verbs occurred in transitive versus intransitive sentences. Nevertheless, it might be that the variable of interest, internal versus external causation, was confounded with some other, unknown, variable.

As mentioned above in the introduction to the experiments, Experiment 1 used a stronger design. First, the sentences were anomalous (cf. McElree & Griffith, 1995, 1998). This reduces considerably the likelihood that some

semantic or associative interaction between the verb and the other words of the sentences could have been confounded with internal versus external causation. Second, because the sentences were anomalous, each sentence frame could be used with both internal and external causation verbs.

Experiments 2, 3, and 4 provide additional data. Experiments 2 and 3 bolstered McKoon and Macfarland's (2000) results with a different paradigm. In McKoon and Macfarland's experiments, sentence processing time was measured as the time required to judge sentence acceptability. It is often thought that acceptability judgments add additional complexity to the cognitive processes used in sentence comprehension. The problem is thought to be that the yes/no decision required by the judgment task invokes decision processes that are not normally part of comprehension, with the consequence that any obtained response time differences might be due to the decision processes, not the comprehension processes of interest.

This potential problem can be addressed in two ways. One is by experimental design. If the judgment task is made as easy as possible, then the contribution of the yes/no decision process to response time can be minimized, thereby reducing the likelihood that it is responsible for obtained response time differences. This was the tack taken by McKoon and Macfarland (2000). The sentences of interest were acceptable, in contrast to the unacceptable filler sentences, which were written to be as unacceptable as possible (e.g., *The toaster allowed*).

Here, we addressed the problem by repeating the experiments without requiring acceptability judgments. In Experiments 2 and 3, a more typical dependent variable was used, sentence reading time. Subjects were asked to read sentences and comprehension was checked with true/false test statements. In Experiment 2, there were intransitive sentences for 14 external causation verbs and 14 internal causation verbs, and in Experiment 3, there were transitive sentences for 14 external and 14 internal causation verbs. For 8 of the verbs of each set, the probability that the verb appeared in a transitive construction in the corpus was low and for 6, it was higher (see Table 2).

6.1. Method

Materials. One acceptable intransitive sentence was constructed for each of the 14 external and 14 internal causation verbs used in the experiment. The mean Kučera–Francis frequencies of the verbs were 6.1 and 6.6 for the internal and external causation verbs, respectively. For the internal causation verbs, the mean probability of appearing in a transitive construction in the corpus is .06 for the low probability verbs and .45 for the higher probability verbs. For the external causation verbs, the corresponding probabilities are .06 and .48. Each sentence contained exactly three words: an article or possessive pronoun, a noun, and the verb, the verb always in past tense.

The numbers of characters per sentence averaged 20.7 for the sentences with internal cause verbs and 20.4 for the sentences with external cause verbs. The frequencies of the nouns of the two sets of sentences were roughly equated (and not significantly different). Also, the two sets of sentences were matched in terms of plausibility, as rated by an independent group of subjects. On a scale of 1 to 5, with 5 being most plausible, the mean for the sentences with internal causation verbs was 4.2 and the mean for the sentences with external causation verbs was 4.1.

The sentences were closely modeled on intransitive uses of the verbs in the corpus. For each verb, the noun that was the subject of the experimental sentence was the kind of entity that was most frequently used as subject for that verb in intransitive constructions in the corpus, except in three cases (two external causation verbs and one internal causation verb) for which the subject was the second most frequent kind of entity. For example, *The wind abated* was taken from corpus sentences about wind abating and *The flowers bloomed* was taken from corpus sentences about flowers blooming. Other examples of sentences with external causation verbs are *The residents awoke*, *The concrete crumbled*, *The signal faded*, and *The crops shriveled*. Other examples of sentences with internal causation verbs are *The roof deteriorated*, *The beams rusted*, *The cider fermented*, and *His knee swelled*. There was one error in creating the sentences; the sentence *The seedlings sprouted* is anomalous because seedlings are plants that have already sprouted (Webster's dictionary). Data for this sentence were eliminated from the analyses. The sentences are given in Appendix B.

There were also filler sentences, including 3-word intransitive sentences and sentences of other structures ranging in length from 5 to 12 words. For each filler sentence, there was a true/false test sentence.

Procedure and subjects. All stimuli were displayed on the screen of a PC monitor. The sentences were presented in short lists, each list followed by three true/false test sentences. At the beginning of each list, subjects pressed the space bar to begin the list. Then the sentences were displayed one at a time. Subjects were to read each sentence and press the space bar when they finished. Then the screen was cleared, there was a 500-ms pause, and the next sentence was displayed. Subjects were instructed to read the sentences carefully enough that they could respond accurately on the true/false tests. An error on a true/false test sentence was followed by the message ERROR for 900 ms.

For 34 subjects, there were 18 lists of 9 sentences per list, preceded by 1 practice list; the experimental sentences for this experiment were mixed among the experimental sentences for another experiment and filler sentences. For the other 28 subjects, there were only the experimental sentences for this experiment and fillers; they were presented in 7 lists of 9 sentences per list. At most, 1 experimental sentence with an internal cause verb and 1 with an external cause verb were presented in each list. The sentences in a list were presented in random order, a new random order after every

second subject. The data from the two groups of subjects were combined for analysis. The subjects were Northwestern University undergraduates participating in the experiment for credit in an Introductory Psychology course.

6.2. Results

Mean reading times were calculated for each subject in each condition, and means of these means are shown in Table 4. The data from four subjects with reading times averaging longer than 3 s for the three-word experimental sentences and the data from three subjects who made more than 25% errors on the true/false tests were eliminated, leaving the 62 subjects described under Section 6.1. Reading times showed considerable variability, and times longer than 2.5 standard deviations above each subject's mean were eliminated (a total over all subjects of about 3% of the data points). The mean proportion correct for true test statements was .92 (1874 ms mean response time) and the mean proportion correct for false test statements was .81 (2159 ms mean response time).

The prediction was that the intransitive sentences with external causation verbs would have longer reading times than the intransitive sentences with internal causation verbs, and the data were consistent with this prediction. The difference in reading times was significant, $F1(1,61) = 10.0$ ($MSE = 35{,}183$) and $F2(1,26) = 4.2$ ($MSE = 12{,}330$). The standard error of the means was 24 ms.

Table 4 shows the reading times for the sentences with external versus internal causation verbs split into sentences with verbs that have a higher probability of occurring in the corpus in transitive constructions and sentences with verbs that have a low probability of occurring as transitives. The first thing to note is that the external–internal difference in reading times appeared for sentences at both levels of transitive probability; the external–internal difference was not significantly different for the two probability conditions, F's $\leqslant 2.7$. Second, although with internal causation verbs, the sentences with higher probability transitive verbs were read more slowly, the overall effect of probability of transitive usage was not significant, F's $\leqslant 2.3$.

For nine of the external causation and nine of the internal causation verbs, their Kučera–Francis frequency was less than 6.0. The difference in

Table 4
Data from Experiment 2

	Externally caused α CAUSE (x (BECOME IN STATE))	Internally caused x (BECOME IN STATE)
	Reading time (in ms)	
All sentences	1664	1557
Low probability transitive	1660	1523
Higher probability transitive	1670	1612

reading times for sentences with external versus internal causation verbs held up for sentences with these low frequency verbs, with reading times of 1700 and 1562 ms, respectively.

To check whether preexperimental associations between the subject nouns and verbs of the sentences could have affected the results, ratings were collected from 12 subjects. For each sentence, we paired the noun and verb (e.g., for the sentence *the flowers bloomed*, the pair was *flowers–bloom*) and asked the subjects to rate how related the two words were on a scale of 1 to 5, with 5 designated most related. Overall, the mean rating for the pairs with internal causation verbs was somewhat higher than that for the pairs with external causation verbs, 3.6 versus 2.8. However, for 3 of the low probability transitive, external causation pairs and 3 of the higher probability transitive, external causation pairs, the mean rating was 3.2, which was about equal to the mean rating, 3.4, for 3 of the low probability transitive and 3 of the higher probability transitive, internal causation pairs. For these 12 sentences, that is, sentences that are equated for preexperimental associations, the external–internal difference in reading times was still large, 1720 ms versus 1654 ms.

7. Experiment 3

The hypothesis guiding these experiments is that the lexical semantic event structure of a verb is a part of the verb's meaning that is used by comprehension processes whenever the verb is encountered, no matter what the syntactic construction in which the verb appears. The external causation verbs have more complex event structures than the internal causation verbs, and so it is predicted that sentences with external causation verbs should have longer comprehension times when the sentences are acceptable transitives as well as when they are acceptable intransitives. In Experiment 3, 14 external causation verbs and 14 internal causation verbs were tested in transitive sentences. As in Experiment 2, 8 of the verbs of each set appeared as transitives in the corpus with low probability and 6 with higher probability.

7.1. Method

Materials. The mean Kučera–Francis frequency for the external causation verbs was 5.4 and the mean for the internal causation verbs was 6.1. The mean probability of appearing in a transitive construction for the eight low probability external causation verbs was .06 and the mean probability for the six higher probability external causation verbs was .44. The corresponding means for the internal causation verbs were .06 and .45.

For each verb, two acceptable transitive sentences were constructed. We used two sentences for each verb to give greater power, hoping to avoid the large amount of variability in reading times observed in Experiment 2. The

sentences all had the structure *noun phrase verb noun phrase*. In all cases, the subject and object noun phrases for each verb were typical of those found with the verb in transitive sentences in the corpus. For example, we found *explode* occurring in a full transitive sentence most often with people exploding bomblike things, so the experimental sentences with *explode* used people as their subjects and nuclear devices and missiles as their objects. *Wilt* usually occurs in transitive sentences with heat or the sun as subject and plants or (groups of) people as object. Accordingly, we constructed experimental sentences with *wilt* that used heat and the sun as subjects and people and plants as objects. Other examples of sentences with external causation verbs are *The nervous teenager crumbled the cigarettes*, *The telephone call awoke the reporter*, *Marriage mellowed the rock star*, and *The sunset reddened the evening sky*. Other examples for internal causation verbs are *The yeast cultures fermented the beer*, *The flood water rotted the carpets*, *The salt water rusted the machinery*, and *Acid rain corroded the building*. The complete set of sentences is shown in Appendix C.

Across all the sentences, the mean number of words per sentence was 5.6 (34.4 characters) for the sentences with external causation verbs and 5.7 (35.5 characters) for the sentences with internal causation verbs. Each sentence with an external causation verb was paired with a sentence with an internal causation verb of about the same Kučera–Francis frequency. The frequency values of the words in the sentences other than the verb were also matched across the two sets of verbs so that they were not significantly different. The two sets of sentences were matched on plausibility, as rated by an independent group of subjects. The mean ratings were 4.4 for the sentences with external causation verbs and 4.3 for the sentences with internal causation verbs.

There was also a pool of 92 filler sentences ranging in length from 3 to 13 words, each with a true/false test question.

Procedure and subjects. The procedure was generally the same as for Experiment 2. The experimental and filler sentences were displayed in 14 lists of 8 sentences per list, each list followed by 3 true/false questions. Each list of 8 sentences was made up of 2 sentences with external causation verbs, 2 with internal causation verbs, and 4 filler sentences. The two sentences that used the same verb could not appear in the same list. Otherwise, the order of the sentences was randomly determined, a new random order for every second subject. Twelve Northwestern undergraduates participated in the experiment for course credit.

7.2. Results

Means were calculated for each subject in each experimental condition, and means of these means are shown in Table 5. For true test statements, the proportion correct was .88 and the mean response time was 2032. For

Table 5
Data from Experiment 3

	Externally caused α CAUSE (x (BECOME IN STATE))	Internally caused x (BECOME IN STATE)
	Reading time (in ms)	
All sentences	2561	2405
Low probability transitive	2516	2390
Higher probability transitive	2519	2318

false test statements, the proportion correct was .76 and the mean response time was 2555 ms.

The main result was that sentences with external causation verbs were read significantly more slowly than sentences with internal causation verbs, $F1(1, 11) = 14.8$ ($MSE = 10{,}980$), and $F2(1, 13) = 7.3$ ($MSE = 23{,}151$) (for the $F2$ ANOVA, the items were paired). The standard error of the means was 30 ms.

Table 5 shows the data broken down by the verbs' low versus higher probabilities of transitive usage. The external–internal difference did not interact with probability level and there was no main effect of probability level (F's < 1.0). For these means (but not the means for all items shown in the first row of the table), the sentences for one pair of verbs (*stagnate*, internal causation, and *proliferate*, external causation) were excluded because their mean reading times were much longer, averaging 3104 ms, than those of any other sentences.

We checked that the external–internal difference held up for verbs with Kučera–Francis frequencies less than the overall mean frequency (less than 6), and it did; the external–internal difference was still large for sentences with these verbs, 2650 ms versus 2441 ms.

To check whether preexperimental associations among the content words of the sentences could have affected our results, ratings were collected for pairs of words from the sentences according to how related they were (on a scale of 1 to 5 where $5 =$ *most related*). For each sentence, there were three pairs to be rated, the subject noun and the verb, the verb and the object noun, and the subject noun and the object noun. Four subjects rated each of the pairs (12 subjects total). The mean rating for the pairs from sentences with external causation verbs was 2.58, and for internal causation verbs it was 2.96. Eliminating three of the internal causation verbs' sentences changed the mean rating for the remainder of the internal causation verb pairs to 2.65, about equal to the rating for the external causation verbs' pairs. Eliminating these three items from the analyses of reading times, the mean for internal causation sentences was 2420 ms, little different from the mean for all the internal causation sentences.

8. Experiment 4

In Experiments 1, 2, and 3, processing times for sentences with externally caused change of state verbs were longer than processing times for sentences with internally caused change of state verbs. We attribute this to the verbs' event templates: the lexical entry for an externally caused change of state verb has a more complex event template than that for an internally caused change of state verb. Across the experiments reported here and in McKoon and Macfarland (2000), we have shown the template complexity effect with both acceptability judgment and reading time procedures, and with intransitive and transitive sentences, as well as anomalous transitive sentences.

An essential and interesting question has not yet been addressed, and that is whether the effect of template complexity appears only in sentence processing, that is, only when a template might be used to guide encoding of a sentence's meaning structure or whether it is such an intrinsic part of processing for verbs that it appears in other tasks. If event templates do reside in verbs' lexical entries and play a part in all verbal processing, then we might expect to see template effects even when a verb is presented as a single word for lexical decision. The differences in response times in Experiments 1, 2, and 3 were reasonably large, averaging about 7% of total response times. Therefore, even though lexical decision is a simple task with short response times, the effect size might be large enough to be observable.

8.1. Method

Materials. The verbs used in the experiment are shown in Table 2. In order to best match verbs of the two types, one verb, *erode*, was used twice in the design of the experiment (but each subject saw it only once in the experiment). The internal and external verbs were paired such that the two members of a pair were as closely matched as possible on the following dimensions. The mean Kučera–Francis frequencies of the verbs were 5.4 and 5.8 for the internally and externally caused verbs, respectively. The two sets of verbs were also matched on imageability: the mean ratings (1 to 5 scale) were 2.4 and 2.5, respectively. Forster (2000) has suggested that words in different conditions in lexical decision experiments should be equated on subjective frequency; that is, they should be equated on subjects' ratings of their frequency of occurrence. For the internally and externally caused verbs, their mean subjective frequency ratings (on a 1 to 5 scale with 5 being *least frequent*) are 3.7 and 3.6, respectively. The two sets of verbs were also matched on the number of other words orthographically similar to them, which was measured as the number of words in the Kučera–Francis norms that differed from a verb by one letter. For the internally and externally caused verbs respectively, the mean numbers of similar words were 2.0 and 1.6. Finally, the number of different senses of meaning per verb, according to WordNet, averaged 1.7 for both sets of verbs. On

none of these dimensions was the difference between verbs of the two classes statistically significant.

Procedure and subjects. Subjects were presented with lists of words and nonwords, displayed one at a time on a PC monitor. Subjects began each list by pressing the space bar on the PC keyboard. For each item, they were instructed to press the "/?" key on the keyboard if it was a word and to press the "z" key if it was not a word. As soon as a key was pressed, the test item disappeared from the screen, there was a 50-ms pause, and then if the response was correct, the next test item was displayed. If a response was not correct, the message "ERROR" was displayed for 900 ms before the 50-ms pause. The subjects were instructed to respond as quickly and accurately as possible.

For 38 subjects, there were 14 lists of test items, preceded by 1 practice list. In each list, there was at most one externally caused and one internally caused change of state verb, each immediately preceded in the test list by a nonword. The remaining items in the list were fillers, 6 words and 8 nonwords. For 40 subjects, there were 11 lists of test items, one externally caused and one internally caused verb in each list (each preceded by a nonword), with 8 word fillers and 10 nonword fillers in each list. The subjects were Northwestern undergraduates participating in the experiment for course credit.

8.2. Results

Mean response times were calculated for each subject and each item in each condition. The mean response time for filler words was 725 ms (.92 correct) and the mean response time for nonwords was 846 ms (.92 correct).

The mean response time for externally caused change of state verbs was 761 ms and the probability of a correct response was .93. Responses for internally caused change of state verbs were faster and more accurate, averaging 728 ms and .97 probability correct. The standard error of the mean was 6 ms. Both the response times and the proportions correct were significantly different for the two classes of verbs, $F1(1,77) = 14.6$ ($MSE = 2,724$) and $F2(1,10) = 7.2$ ($MSE = 1,933$) for response times and $F1(1,77) = 10.0$ ($MSE = .005$) and $F2(1,10) = 3.8$ ($MSE = .002$; $p = .07$) for proportion correct. We checked that the differences held up for low frequency verbs (frequencies less than 6), and they did: for externally caused verbs, 802 ms and .92 probability correct; for internally caused verbs, 756 ms and .97 probability correct.

9. General discussion

The data from the four experiments presented here and those presented by McKoon and Macfarland (2000) display a remarkable uniformity: Sen-

tences with externally caused change of state verbs have longer processing times than sentences with internally caused change of state verbs. What makes this result especially compelling is that it holds for infrequently used verbs as well as frequently used ones, it holds for anomalous sentences as well as acceptable sentences, it holds for transitive sentences as well as intransitive sentences, it holds for sentences that match the most frequent constructions in which a verb is used as well as for sentences that do not, and it holds for acceptability judgment times as well as reading times. In addition, lexical decision response times are slower for externally caused change of state verbs than internally caused change of state verbs.

From the MTS view, the processing time difference comes about because internally caused change of state verbs have a different, simpler, event template than externally caused change of state verbs. The event templates capture only the syntactically relevant parts of verbs' meanings, not all the other parts of verbs' meanings, yet the differences in template complexity had a sizeable effect on processing times for the verbs themselves in lexical decision, about 30 ms, and on processing times for sentences in which the verbs occurred, about 100 ms.

The empirical diagnostic by which a verb is classified as denoting an internally or externally caused change of state is that, for transitive sentences with concrete objects, the range of entities appearing in subject position is more constrained for the internal than the external causation verbs. The reason for this constraint is that the entity in subject position must be a participant in the single, internally controlled, change of state event for an internal causation verb, but for an external causation verb, the entity in subject position is not a participant in the change of state event; it is a participant in a separate, causing subevent.

The diagnostic can be observed only when the verbs appear in transitive sentences. This means that it is difficult to conceive of any direct relation between the statistics of the diagnostic and the results of the processing time experiments. The diagnostic does not apply to intransitive sentences, yet the processing time advantage for the internal causation verbs shows up for intransitive sentences as well as for transitive sentences. Some verbs appear very rarely in transitive sentences, and therefore the diagnostic rarely applies to them, yet the processing time advantage shows up for them as well as for verbs which frequently appear in transitive sentences. And the processing time advantage shows up in lexical decision, where there is no sentence at all. Instead of relating the processing time difference directly to the diagnostic, we explain both with the theoretical construct of event templates: a simpler, internal causality template for the internally caused change of state verbs and a more complex, external causality template for the externally caused change of state verbs.

A critical element of our explanation is that, in MTS, there is only a single event template underlying all uses of a verb. If instead, the lexical entry

for a verb included a separate representation of structural information for each possible use of the verb, then it would be difficult to explain the data obtained in the experiments. If both the internal and the external causation verbs had lexical representations of both transitive and intransitive structures, then there would be no reason to expect shorter response times with the internal causation verbs.

According to the MTS view, when an event is denoted by an externally caused change of state verb (α CAUSE (x (BECOME IN STATE))), it is conceptualized as involving both an external causing event and a change of state event. This is true even when the verb is used in an intransitive construction and the causing event is not explicitly stated. When an event is denoted by an internally caused change of state verb (x (BECOME IN STATE)), it is conceptualized as involving only a change of state. This is true even when a second participant is mentioned as the subject of the verb in a transitive sentence. The second participant is understood as being a part of the single, change of state, event.

Exactly how the difference in the internal versus external causality templates affects sentence comprehension is something about which we can speculate. The extra time required for understanding an event denoted by an external causation verb could come about for any or all of several reasons: The externally caused templates have two subevents and the internally caused templates only one; the externally caused but not the internally caused templates include the CAUSE relation; the externally caused templates offer two possible subevent assignments for arguments and the internally caused templates only one; the externally caused but not the internally caused templates embed the x (BECOME IN STATE) template in a larger template structure; the causing subevent in the externally caused template can sometimes be represented with one of several possible arguments, whereas the internal causation template has no such ambiguity; for an external causation verb, the cause can be a whole event (α) but for an internal causation verb, it can be only a single entity; and so on. At this point, there are no data to distinguish among the possibilities.

9.1. Other approaches

The event template approach, and predicate decompositions in general, are only just beginning their theoretical development. Therefore it is important to consider possible alternative explanations of the empirical differences we have observed between internally and externally caused change of state verbs. For example, it might be thought possible that the data could be explained by some syntactic difference between verbs of the two classes. We can rule out a difference based on the probabilities of transitive versus intransitive occurrences, but we cannot absolutely rule out some other, as yet unproposed, syntactic difference. However, it seems unlikely that a syn-

tactic hypothesis could explain the sentence production data from the corpus because the constraint shown by the corpus data is a semantic constraint: the subjects of transitive sentences with internally caused change of state verbs are statistically constrained to have subjects that are inherent participants in the change of state events they denote. It would be difficult to explain this constraint in syntactic terms.

Besides event templates, there are a number of other formulations of semantic predicate decompositions, and our findings are consistent with some of these (e.g., Goldberg, 1995; Jackendoff, 1996). Jackendoff, for example, represents the events denoted by verbs as relations among entities, with the relations evolving over time. For internally caused changes of state, the single subevent could evolve through time as the changing entity progressed into its new state. If, as with the verbs we studied, a natural force participant was inherently involved in the change of state process, then that participant would be bound through time to the progress of the change of state. For externally caused changes of state, the causing subevent could occur and evolve through time independently of the change of state subevent (except that the causing subevent could not begin later than the change of state subevent). With these assumptions, the representation of an externally caused event would be more complex than the representation of an internally caused event, and so consistent with our data.

A crucial aspect of the findings presented here is that the external–internal difference appeared with both transitive and intransitive sentences. This indicates to us that in the entry in the lexicon for a verb, there is only a single lexical semantic structure to represent syntactically relevant information (unless the verb is truly ambiguous). A single representation of syntactically relevant information runs counter to some claims in the lexical semantics literature. Pustejovsky (1991) (but see Pustejovsky, 1995) and Van Valin and LaPolla (1997) have proposed that change of state verbs have different representations in their transitive and intransitive uses. For *John closed the door*, the lexical entry for *close* would include both a causing subevent and a change of state subevent, but for *the door closed*, there would be only a change of state event. By this view, the intransitive *close* would have the same lexical semantic representation as internally caused change of state verbs and so, for intransitive sentences, no difference between them in processing times would be predicted, contrary to the data.

In some current psycholinguistic theories, syntactic processing occurs in an encapsulated module separate from other cognitive systems and operates on sentences prior to the other systems (Frazier, 1987; Frazier & Clifton, 1996; Frazier & Rayner, 1982; Rayner, Carlson, & Frazier, 1983). These theories could be compatible with our data if the syntactic module parses sentences into syntactic structures which are then forwarded to semantic processes that make use of the event template information postulated by MTS. Under this scenario, transitive and intransitive sentences could be

translated into their appropriate syntactic structures with no processing time difference between externally and internally caused change of state verbs. Processing time differences would occur once the syntactic structures were transferred to semantic processing systems. We know of no empirical evidence that would either confirm or deny this scenario.

9.2. Constraint-based models

There are several ways the MTS view differs from most constraint-based models. While frequency information plays no role in MTS, the central assumption underlying constraint-based models is that processing is easier and faster for linguistic structures that occur more frequently than for structures that occur less frequently (e.g., MacDonald et al., 1994; Trueswell, Tanenhaus, & Garnsey, 1994). It is not always specified exactly what kinds of linguistic information are encoded for each word in the lexicon, but we thought it reasonable that the encodings of verbs might include the probability with which they appear in transitive versus intransitive constructions. For this reason, in each experiment, we varied the mean probability of transitive usage from about .06 for one group of verbs to about .46 for the other group. This large difference in probability had no significant effect on sentence processing times. This null finding does not, of course, refute the models. It could be that transitive/intransitive probabilities are not encoded with verbs, or perhaps they are encoded but either the manipulation of probability was not large enough or the effect was so small as to be empirically unobservable.

A second way in which MTS and some constraint-based models (e.g., Boland, 1997; MacDonald et al., 1994; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Tanenhaus, Spivey-Knowlton, & Hanna, 2000; Trueswell et al., 1994) differ is that, for the models to encode the relative probabilities of occurrence for each of several different structures in which a verb can occur, the models must have a separate representation of each of the structures. There must be separate representations to which the different probabilities can be attached. In MTS, there is only a single representation of syntactically relevant information for each verb in the lexicon (unless the verb is truly ambiguous), so there is no way to encode different probabilities for, for example, active versus passive structures or transitive versus intransitive structures.

A number of constraint-based models make specific proposals about the kinds of information encoded lexically for verbs. Boland (1997) has proposed that verbal entries include both thematic role lists, representing semantic information, and subcategorization frames, representing syntactic information. MacDonald et al. (1994) (see also MacDonald, 1994) has suggested that a verb's lexical entry includes thematic role lists for semantic information (p. 683) and X-bar structures for syntactic information (p. 684).

In the models proposed by McRae et al. (1998), Spivey and Tanenhaus (1998), Tanenhaus et al. (2000), Trueswell et al. (1994), and Trueswell (1996), semantic information is specified in thematic role lists, and syntactic information is specified in the probabilities with which a verb appears in active versus passive constructions (because the models were designed to explain the comprehension of main clauses versus reduced relative clauses). Garnsey, Pearlmutter, Myers, and Lotocky (1997) suggest that lexical entries contain information about whether a verb can appear in constructions with a direct object (syntactic information) and about a verb's thematic role lists. McElree and Griffith (1995, 1998) summarize empirical results that have been taken by various theorists as support for the inclusion of both thematic role lists and subcategorization frames in verbal lexical entries. All of the approaches just listed contrast with the MTS framework in that, in MTS, the lexical entry for a verb contains no verb-specific syntactic information.

For all of these models, it is not clear how they would handle the results of the experiments reported here. One possibility is the same as that suggested for the modularity models (e.g., Frazier, 1987). Transitive and intransitive sentences are translated into their appropriate syntactic structures with no processing time difference between externally and internally caused change of state verbs, and the external–internal difference occurs only at some later stage of processing. Whether there are other ways the models can handle the data presented here is an open question. In general, it is not possible to evaluate the models because they have not yet been developed to the point of making explicit assumptions about representations of lexical information of the kind needed here.

10. Summary

The heart of the MTS view is that syntax is both a reflection of linguistic meaning and a fruitful way to explore linguistic meaning. For the verbs of interest in this article, we examined large numbers of naturally produced sentences in order to understand the syntactic constructions in which the verbs are used (McKoon & Macfarland, 2000). We were able to relate summary statistics from the corpus sentences directly to the verbs' meanings, that is, to the syntactically relevant portions of the verbs' meanings that are expressed in event templates. The results of the sentence comprehension and lexical decision experiments converge with the corpus statistics to provide significant support for the psycholinguistic importance of event templates.

The MTS view outlines a research agenda for which the goal is to bring together linguistic meaning and the behaviors of linguistic elements, as reflected in sentence structures and in sentence comprehension. Externally

and internally caused change of state verbs provide an especially useful case in point. The constraints on the behaviors of the verbs in sentences are subtle, yet observable in the statistics of natural usage. The same theoretical constructs, event templates, explain both the sentence behavior and predict verb and sentence comprehension times.

It is also worth noting that event templates are, in essence, an hypothesis about the decompositionality of verb meanings. The syntactically relevant part of a verb's meaning is decomposed into participants in the events and subevents described by the verb and relations among those participants. In supporting this notion, the data presented in this article refute Fodor's often repeated claim that there is no evidence that the meanings of words are psycholinguistically decomposable (e.g., Fodor, Fodor, & Garrett, 1975; Fodor, Garrett, Walker, & Parkes, 1980; Fodor & Lepore, 1998; see also Cutler, 1983; Kintsch, 1974; Rayner & Duffy, 1986).

In the experiments described here, we have provided only limited initial support for the MTS view because we examined only a few verbs from only two verb classes. Even so, the empirical work was comprehensive, including three different experimental paradigms with three different types of sentences plus lexical decision on the verbs themselves, as well as exhaustive corpus studies of naturally produced sentences. Although the scope was limited to only the two classes of verbs, we point to the success of the hypotheses in accounting for the data as a beginning in the MTS approach to how lexical semantic information plays a role in sentence processing and how considerations of meaning, in general, can come to be incorporated into sentence processing research. As more and more verb classes and syntactic phenomena are examined, we expect a more comprehensive picture to emerge about how sentence structures are connected, through meaning, to language processing.

Appendix A. Unacceptable transitive sentences for Experiment 1

Each sentence frame is shown with the four verbs with which it was used. In each case, the first and third verbs are externally caused, the second and fourth internally caused.

Some painters abate/germinate/atrophy/stagnate medals.
Some athletes abate/germinate/atrophy/stagnate brushes.
Some hunters abate/germinate/atrophy/stagnate clocks.
Some tourists abate/germinate/atrophy/stagnate needles.

Some candidates crumble/deteriorate/explode/wither victories.
Some teams crumble/deteriorate/explode/wither votes.
Some babies crumble/deteriorate/explode/wither shadows.
Some owners crumble/deteriorate/explode/wither clouds.

Some editors shrivel/rot/vibrate/wilt oceans.
Some captains shrivel/rot/vibrate/wilt mistakes.
Some readers shrivel/rot/vibrate/wilt weddings.
Some ministers shrivel/rot/vibrate/wilt submarines.

Some criminals dissipate/sprout/fray/ferment boats.
Some fishermen dissipate/sprout/fray/ferment prisons.
Some immigrants dissipate/sprout/fray/ferment kitchens.
Some dentists dissipate/sprout/fray/ferment loans.

Some actors oxidize/corrode/redden/erode lunches.
Some visitors oxidize/corrode/redden/erode scripts.
Some males oxidize/corrode/redden/erode airports.
Some historians oxidize/corrode/redden/erode pensions.

Some farmers fossilize/blister/thaw/swell clues.
Some detectives fossilize/blister/thaw/swell barns.
Some engineers fossilize/blister/thaw/swell taxes.
Some manufacturers fossilize/blister/thaw/swell cabins.

Appendix B. Acceptable intransitive sentences for Experiment 2

Externally caused verbs
The wind abated.
His skills atrophied.
The residents awoke.
The concrete crumbled.
The missile exploded.
The signal faded.
The crops shriveled.
The fan vibrated.
The bullet splintered.
The smoke dissipated.
The bones fossilized.
The string frayed.
His skin reddened.
The turkey thawed.

Internally caused verbs
The flowers bloomed.
The roof deteriorated.
The seeds germinated.
The potatoes rotted.
The beams rusted.
The economy stagnated.
The leaves wilted.
The roses withered.
Her skin blistered.
The metal corroded.
The beaches eroded.
The cider fermented.
His knee swelled.

Appendix C. Acceptable transitive sentences for Experiment 3

For each pair of verbs, the sentences with internally caused verbs are listed first, the sentences with externally caused verbs second.

The plants bloomed yellow blossoms.
The trees bloomed pink flowers.
The illness atrophied the lower leg.
The stroke atrophied the right brain.

The heavy traffic deteriorated the bridge.
The harsh climate deteriorated the roads.
The opera singer vibrated the notes.
The famous violinist vibrated the strings.

Communism stagnated the country's economy.
The regulations stagnated private investments.
The senators proliferated government subsidies.
The politicians proliferated welfare programs.

The local florist germinated the seeds.
The amateur botanist germinated the corn.
The French chef crumbled the cheese.
The nervous teenager crumbled the cigarettes.

The intense heat wilted the crowd.
The bright sun wilted the roses.
The fire alarm awoke the residents.
The telephone call awoke the reporter.

The rare disease rotted the potatoes.
The flood water rotted the carpets.
The scientists exploded the nuclear device.
Religious extremists exploded the bomb.

The severe drought withered the crops.
The late frost withered the tulips.
The police abated violent crime.
The doctors abated infectious diseases.

The constant rain rusted the car.
The salt water rusted the machinery.
Age mellowed the rebel leader.
Marriage mellowed the rock star.

The intense sun blistered the paint.
The hot sauce blistered his tongue.
The wind reddened the baby's face.
The sunset reddened the evening sky.

The plants sprouted tender buds.
The trees sprouted tiny leaves.
The solution oxidized the scrap metal.
The water oxidized the iron beam.

Appendix C (*continued*)

The violent storms swelled the sea.
The rainy season swelled the rivers.
The huge radiators dissipated the heat.
The strong winds dissipated the clouds.

The yeast cultures fermented the beer.
The apple growers fermented the cider.
The scandal frayed party unity.
The lengthy strike frayed tempers.

Acid rain corroded the building.
The detergent corroded the fine silver.
The cook thawed the holiday turkey.
The mild temperatures thawed the soil.

The storm eroded the beach.
The river eroded the rocks.
The weather stiffened his joints.
The starch stiffened his shirts.

References

Anderson, S. R. (1971). On the role of the deep structure in semantic interpretation. *Foundations of Language, 6*, 197–219.

Atkins, B. T. S., & Levin, B. (1995). Building on a corpus: A linguistic and lexicographical look at some near synonyms. *International Journal of Lexicography, 8*, 85–114.

Bates, E., & Goodman, J. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes, 12*, 507–584.

Boland, J. E. (1997). The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes, 12*, 423–484.

Burzio, L. (1986). *Italian syntax: A government-binding approach*. Dordrecht: Reidel.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin and use*. New York: Praeger.

Croft, W. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.

Croft, W. (1998). Event structure in argument linking. In M. Butt & W. Geuder (Eds.), *The projection of arguments: Lexical and compositional factors* (pp. 21–64). Stanford, CA: CSLI.

Cutler, A. (1983). Lexical complexity and sentence processing. In G. B. Flores d' Arcais & R. J. Jarvella (Eds.), *The processes of language understanding*. New York: Wiley.

Dowty, D. R. (1979). *Word meaning and Montague grammar*. Dordrecht: Reidel.

Dowty, D. R. (1989). On the semantic content of the notion 'thematic role.' In B. Partee, B. Chierchia, & R. Turner (Eds.), *Properties types and meanings* (Vol. II) (pp. 69–130). Dordrecht: Kluwer.

Dowty, D. R. (1991). Thematic proto-roles and argument selection. *Language, 67*, 547–619.

Fellbaum, C. (Ed.). (1998). *WordNet*. Cambridge, MA: MIT Press.

Fodor, J. D., Fodor, J. A., & Garrett, M. F. (1975). The psychological unreality of semantic representations. *Linguistic Inquiry, 6*, 515–531.

Fodor, J. A., Garrett, M. F., Walker, E. C. T., & Parkes, C. H. (1980). Against definitions. *Cognition, 8*, 263–367.

Fodor, J. A., & Lepore, E. (1998). The emptiness of the lexicon: Reflections on James Pustejovsky's *The Generative Lexicon. Linguistic Inquiry, 29*, 269–288.

Forster, K. I. (2000). The potential for experimenter bias effects in word recognition experiments. *Memory and Cognition, 28*, 1109–1115.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton–Mifflin.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. London: Erlbaum.

Frazier, L., & Clifton, C. (1996). *Construal*. Cambridge, MA: MIT Press.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178–210.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language, 37*, 58–93.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Grimshaw, J. (1990). *Argument structure*. Cambridge, MA: MIT Press.

Grimshaw, J. (1993). *Semantic structure and semantic content*. Unpublished manuscript, Rutgers University, New Brunswick, NJ.

Hale, K., & Keyser, S. J. (1993). On argument structure and the lexical expression of syntactic relations. In K. Hale, & S. J. Keyser (Eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (pp. 53–109). Cambridge, MA: MIT Press.

Hay, J., Kennedy, C., & Levin, B. (1999). Scalar structure underlies telicity in degree achievements. In *Proceedings of SALT 9* (pp. 127–144). Ithaca, NY: Cornell Linguistics Circle Publications, Cornell University Press.

Jackendoff, R. S. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jackendoff, R. S. (1996). The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language and Linguistic Theory, 14*, 305–354.

Kiparsky, P. (1985). *Morphology and grammatical relations*. Unpublished manuscript, Stanford University, Stanford, CA.

Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.

Levin, B., & Rappaport Hovav, M. (1995). *Unaccusativity: At the syntax–lexical semantics interface*. Cambridge, MA: MIT Press.

Levin, B., & Rappaport Hovav, M. (1996). *From lexical semantics to argument realization*. Unpublished manuscript, Northwestern University and Bar Ilan University, Evanston, IL and Ramat Gan, Israel.

MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes, 9*, 157–201.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676–703.

Marantz, A. P. (1984). *On the nature of grammatical relations*. Cambridge, MA: MIT Press.

McElree, B., & Griffith, T. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 134–157.

McElree, B., & Griffith, T. (1998). Structural and lexical constraints on filling gaps during sentence processing: A time-course analysis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*, 432–460.

McKoon, G., & Macfarland, T. (2000). Externally and internally caused change of state verbs. *Language, 76*, 833–858.

McKoon, G., & Ratcliff, R. (submitted for publication). *Meaning through syntax*: *Language comprehension and the reduced relative clause construction*.

McKoon, G., Ratcliff, R., Ward, G., & Sproat, R. (1993). Syntactic prominence effects on discourse processes. *Journal of Memory and Language, 32*, 593–607.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 38*, 283–312.

Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics, 14*, 15–28.

Perlmutter, D. M. (1978). Impersonal passives and the unaccusative hypothesis. In *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society* (pp. 157–189). Berkeley: University of California, Berkeley Linguistics Society.

Pustejovsky, J. (1991). The syntax of event structure. *Cognition, 41*, 47–81.

Pustejovsky, J. (1995). *The generative lexicon: A computational theory of lexical semantics*. Cambridge, MA: Bradford Books/MIT Press.

Rappaport, M., & Levin, B. (1988). What to do with theta-roles. In W. Wilkins (Ed.), *Syntax and semantics 21: Thematic relations* (pp. 7–36). New York: Academic Press.

Rappaport Hovav, M., & Levin, B. (1998). Building verb meanings. In M. Butt & W. Geuder (Eds.), *The projection of arguments: Lexical and compositional factors* (pp. 97–134). Stanford, CA: CSLI.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior, 22*, 358–374.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition, 14*, 191–201.

Smith, C. S. (1970). Jespersen's 'Move and Change' class and causative verbs in English. In M. A. Jazayery, E. C. Polome, & W. Winter (Eds.), *Linguistic and literary studies in honor of Archibald A. Hill: Vol. 2. Descriptive linguistics* (pp. 101–109). The Hague: Mouton.

Smith, C. S. (1991). *The parameter of aspect*. Dordrecht: Kluwer.

Song, G. (1996). Causation, adicity and lexical aspect. In M. Przezdziecki & L. Waley (Eds.), *Proceedings of ESCOL '95* (pp. 299–307).

Speas, M. J. (1990). *Phrase structure in natural language*. Dordrecht: Kluwer.

Spivey, M., & Tanenhaus, M. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology, 24*, 1521–1543.

Stevenson, S., & Merlo, P. (1997). Lexical structure and parsing complexity. *Language and Cognitive Processes, 12*, 349–399.

Tanenhaus, M., Spivey-Knowlton, M., & Hanna, J. (2000). Modeling thematic and discourse context effects with a multiple constraints approach: Implications for the architecture of the language comprehension system. In M. Crocker, M. Pickering, & C. Clifton (Eds.), *Architectures and mechanisms of language acquisition and processing* (pp. 90–118). Cambridge, UK: Cambridge University Press.

Tenny, C. L. (1994). *Aspectual roles and the syntax–semantics interface*. Dordrecht: Kluwer.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language, 35*, 566–585.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language, 33*, 285–318.

Van Valin, R. D., Jr., & LaPolla, R. J. (1997). *Syntax: Structure, meaning and function*. Cambridge, UK: Cambridge University Press.

Vendler, Z. (1957). Verbs and time. *Philosophical Review, 56*, 143–160 (also in Vendler, Z. (1967). *Linguistics in philosophy* (97–121). Ithaca, NY: Cornell University Press).
Wasow, T. (1985). Postscript. In P. Sells (Ed.), *Lectures on contemporary syntactic theories* (pp. 193–205). Stanford, CA: CSLI.
Wechsler, S. (1995). *The semantic basis of argument structure*. Stanford, CA: CSLI.
Williams, E. (1981). Argument structure and morphology. *The Linguistic Review, 1*, 81–114.