# Aspectual Modifications to a LCS Database for NLP Applications*

**Bonnie J. Dorr** and **Mari Broman Olsen**
Laboratory for Language and Media Processing
Department of Computer Science and
Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742, USA
{bonnie,molsen}@umiacs.umd.edu

**Abstract:** Verbal and compositional lexical aspect provide the underlying temporal structure of events. Knowledge of lexical aspect, e.g., (a)telicity, is therefore required for interpreting event sequences in discourse (Dowty, 1986; Moens and Steedman, 1988; Passoneau, 1988), interfacing to temporal databases (Androutsopoulos, 1996), processing temporal modifiers (Antonisse, 1994), describing allowable alternations and their semantic effects (Resnik, 1996; Tenny, 1994), and selecting tense and lexical items for natural language generation ((Dorr and Olsen, 1996; Klavans and Chodorow, 1992), cf. (Slobin and Bocaz, 1988)). We show that it is possible to represent lexical aspect—both verbal and compositional—on a large scale, using Lexical Conceptual Structure (LCS) representations of verbs in the classes cataloged by Levin (1993). We show how proper consideration of these universal pieces of verb meaning may be used to refine lexical representations and derive a range of meanings from combinations of LCS representations. A single algorithm may therefore be used to determine lexical aspect classes and features at both verbal and sentence levels. Finally, we illustrate how knowledge of lexical aspect facilitates the interpretation of events in NLP applications.

*This work was supported, in part, by Department of Defense contract MDA90496C1250, Army Research Office contract DAAL03-91-C-0034 through Battelle Corporation, NSF NYI IRI-9357731 and Logos Corporation, NSF CNRS INT-9314583, Advanced Research Projects Agency and ONR contract N00014-92-J-1929, Alfred P. Sloan Research Fellow Award BR3336, Army Research Institute contract MDA-903-92-R-0035 and Microelectronics and Design, Inc., and the University of Maryland General Research Board. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, Department of Army, or Department of Defense.

# 1 Introduction

Knowledge of lexical aspect—how verbs denote situations as developing or holding in time—is required for interpreting event sequences in discourse (Dowty, 1986; Moens and Steedman, 1988; Passoneau, 1988), interfacing to temporal databases (Androutsopoulos, 1996), processing temporal modifiers (Antonisse, 1994), describing allowable alternations and their semantic effects (Resnik, 1996; Tenny, 1994), and for selecting tense and lexical items for natural language generation ((Dorr and Olsen, 1996; Klavans and Chodorow, 1992), cf. (Slobin and Bocaz, 1988)). In addition, preliminary pyscholinguistic experiments (Antonisse, 1994) indicate that subjects are sensitive to the presence or absence of aspectual features when processing temporal modifiers. Resnik (1996) showed that the strength of distributionally derived selectional constraints helps predict whether verbs can participate in a class of diathesis alternations, with aspectual properties of verbs clearly influencing the alternations of interest. He also points out that these properties are difficult to obtain directly from corpora.

The ability to determine lexical aspect, on a large scale and in the sentential context, therefore yields an important source of constraints for corpus analysis and psycholinguistic experimentation, as well as for NLP applications such as machine translation (Dorr et al., 1995) and foreign language tutoring (Dorr et al., 1995; Sams, 1995; Weinberg et al., 1995). Other researchers have proposed corpus-based approaches to acquiring lexical aspect information with varying data coverage: Klavans and Chodorow (1992) focus on the event-state distinction in verbs and predicates; Light (1996) considers the aspectual properties of verbs and affixes; and McKeown and Siegel (1996) describe an algorithm for classifying sentences according to lexical aspect properties. Conversely, a number of works in the linguistics literature have proposed lexical semantic templates for representing the aspectual properties of verbs (Dowty, 1979; Hovav and Levin, 1995; Levin and Rappaport Hovav, To appear), although these have not been implemented and tested on a large scale.

We show that it is possible to represent the lexical aspect both of verbs alone and in sentential contexts using Lexical Conceptual Structure (LCS) representations of verbs in the classes cataloged by Levin (1993). We show how proper consideration of these universal pieces of verb meaning may be used to refine lexical representations and derive a range of meanings from combinations of LCS representations. A single algorithm may therefore be used to determine lexical aspect classes and features at both verbal and sentential levels. Finally, we illustrate how access to lexical aspect facilitates lexical selection and the interpretation of events in machine translation and foreign language tutoring applications, respectively.

| Aspectual Class | Telic | Dynamic | Durative | Examples |
|---|---|---|---|---|
| State | - | - | + | *know, have* |
| Activity | - | + | + | *march, paint* |
| Accomplishment | + | + | + | *destroy* |
| Achievement | + | + | - | *notice, win* |

Table 1: Featural Identification of Aspectual Classes

## 2 Lexical Aspect

Following Olsen (To appear), we distinguish between lexical and grammatical aspect, roughly the situation and viewpoint aspect of Smith (1991). Lexical aspect refers to the type of situation denoted by the verb, alone or combined with other sentential constituents. Grammatical aspect takes these situation types and presents them as imperfective *(John was winning the race/loving his job)* or perfective *(John had won/loved his job)*. Verbs are assigned to lexical aspect classes, as in Table 1 (cf. (Brinton, 1988)[p. 57], (Smith, 1991)) based on their behavior in a variety of syntactic and semantic frames that focus on their features.[1]

A major source of the difficulty in assigning lexical aspect features to verbs is the ability of verbs to appear in sentences denoting situations of multiple aspectual types. Such cases arise, e.g., in the context of foreign language tutoring (Dorr et al., 1995; Sams, 1995; Weinberg et al., 1995), where a a 'bounded' interpretation for an atelic verb, e.g., *march*, may be introduced by a path PP *to the bridge* or *across the field* or by a NP *the length of the field*:

(1) The soldier marched to the bridge.
The soldier marched across the field.
The soldier marched the length of the field.

Some have proposed, in fact, that aspectual classes are gradient categories (Klavans and Chodorow, 1992), or that aspect should be evaluated only at the clausal or sentential level (esp. (Verkuyl, 1993); see (Klavans and Chodorow, 1992) for NLP applications).

Olsen (To appear) showed that, although sentential and pragmatic context influence aspectual interpretation, input to the context is constrained in large part by verbs' aspectual

[1]Two additional categories are identified by Olsen (To appear): Semelfactives (*cough, tap*) and Stage-level states (*be pregnant*). Since they are not assigned templates by either Dowty (1979) or Levin and Rappaport Hovav (To appear), we do not discuss them in this paper.

| **Aspectual Class** | **Telic** | **Dynamic** | **Durative** | **Examples** |
|---|---|---|---|---|
| State | | | + | *know, have* |
| Activity | | + | + | *march, paint* |
| Accomplishment | + | + | + | *destroy* |
| Achievement | + | + | | *notice, win* |

Table 2: Privative Featural Identification of Aspectual Classes

information. In particular, she showed that the positively marked features did not vary: [+telic] verbs such as *win* were always bounded, for example, In contrast, the negatively marked features could be changed by other sentence constituents or pragmatic context: [-telic] verbs like *march* could therefore be made [+telic]. Similarly, stative verbs appeared with event interpretations, and punctiliar events as durative. Olsen therefore proposed that aspectual interpretation be derived through monotonic composition of marked privative features [+/$\emptyset$ dynamic], [+/$\emptyset$ durative] and [+/$\emptyset$ telic], as shown in Table 2 (Olsen, To appear, pp. 32–33).

With privative features, other sentential constituents can add to features provided by the verb but not remove them. On this analysis, the activity features of *march* ([+durative, +dynamic]) propagate to the sentences in (1), with [+telic] added by the NP or PP, yielding an accomplishment interpretation. The feature specification of this compositionally derived accomplishment is therefore identical to that of a sentence containing a telic accomplishment verb, such as *produce* in (2).

(2) The commander produced the campaign plan.

Dowty (1979) explored the possibility that aspectual features in fact constrained possible units of meaning and ways in which they combine. In this spirit, Levin and Rappaport Hovav (To appear) demonstrate that limiting composition to aspectually described structures is an important part of an account of how verbal meanings are built up, and what semantic and syntactic combinations are possible.

We draw upon these insights in revising our LCS lexicon in order to encode the aspectual features of verbs. In the next section we describe the LCS representation used in a database of 9000 verbs in 191 major classes. We then describe the relationship of aspectual features to this representation and demonstrate that it is possible to determine aspectual features from LCS structures, with minimal modification. We demonstrate composition of the LCS and corresponding aspectual structures, by using examples from NLP applications that employ the LCS database.

## 3 Lexical Conceptual Structures

We adopt the hypothesis explored in Dorr and Olsen (1996) (cf. (Tenny, 1994)), that lexical aspect features are abstractions over other aspects of verb semantics, such as those reflected in the verb classes in Levin (1993). Specifically we show that a privative model of aspect provides an appropriate diagnostic for revising lexical representations: aspectual interpretations that arise only in the presence of other constituents may be removed from the lexicon and derived compositionally. Our modified LCS lexicon then allows aspect features to be determined algorithmically both from the verbal lexicon and from composed structures built from verbs and other sentence constituents, using uniform processes and representations.

This project on representing aspectual structure builds on previous work, in which verbs were grouped automatically into Levin's semantic classes (Dorr and Jones, 1996; Dorr, To appear) and assigned LCS templates from a database built as Lisp-like structures (Dorr, 1997). The assignment of aspectual features to the classes in Levin was done by hand inspection of the semantic effect of the alternations described in Part I of Levin (Olsen, 1996), with automatic coindexing to the verb classes (see (Dorr and Olsen, 1996)). Although a number of Levin's verb classes were aspectually uniform, many required subdivisions by aspectual class; most of these divided atelic "manner" verbs from telic "result" verbs, a fundamental linguistic distinction (cf. (Levin and Rappaport Hovav, To appear) and references therein). Examples are discussed below.

Following Grimshaw (1993) Pinker (1989) and others, we distinguish between semantic structure and semantic content. Semantic structure is built up from linguistically relevant and universally accessible elements of verb meaning. Borrowing from Jackendoff (1990), we assume semantic structure to conform to wellformedness conditions based on Event and State *types*, further specialized into *primitives* such as GO, STAY, BE, GO-EXT, and ORIENT. We use Jackendoff's notion of *field*, which carries Loc(ational) semantic primitives into non-spatial domains such as Poss(essional), Temp(oral), Ident(ificational), Circ(umstantial), and Exist(ential). We adopt a new primitive, ACT, to characterize certain *activities* (such as *march*) which are not adequately distinguished from other event types by Jackendoff's GO primitive.[2] Finally, we add a manner component, to distinguish among verbs in a class, such the motion verbs *run*, *walk*, and *march*. Consider *march*, one of Levin's *Run Verbs* (51.3.2);[3] we assign it the template in (3)(i), with the corresponding Lisp format shown in (3)(ii):

[2] Jackendoff (1990) augments the *thematic* tier of Jackendoff (1983) with an *action* tier, which serves to characterize activities using additional machinery. We choose to simplify this characterization by using the ACT primitive rather than introducing yet another level of representation.

[3] The numbers after the verb examples are verb class sections in Levin (1993).

(3) (i) $[_{\text{Event}}\ \text{ACT}_{\text{Loc}}$
$([_{\text{Thing}} * 1],[_{\text{Manner}}\ \text{BY MARCH 26}])]$

(ii)
```
(act loc
  (* thing 1) (by march 26))
```

This list structure recursively associates arguments with their logical heads, represented as primitive/field combinations, e.g., $\text{ACT}_{\text{Loc}}$ becomes `(act loc ...)` with a `(thing 1)` argument. Semantic content is represented by a constant in a semantic structure position, indicating the linguistically inert and non-universal aspects of verb meaning (cf. (Grimshaw, 1993; Pinker, 1989; Levin and Rappaport Hovav, To appear)), the manner component `by march` in this case. The numbers in the lexical entry are codes that map between LCS positions and their corresponding thematic roles (e.g., 1 = *agent*). The $*$ marker indicates a variable position (i.e., a non-constant) that is potentially filled through composition with other constituents.

In (3), `(thing 1)` is the only argument. However, other arguments may be instantiated compositionally by the end-NLP application, as in (4) below, for the sentence *The soldier marched to the bridge*:

(4) (i)
$$
\begin{array}{l}
[_{\text{Event}}\ \text{CAUSE} \\
\quad ([_{\text{Event}}\ \text{ACT}_{\text{Loc}} \\
\quad\quad ([_{\text{Thing}}\ \text{SOLDIER}], \\
\quad\quad\quad [_{\text{Manner}}\ \text{BY MARCH}])], \\
\quad [_{\text{Path}}\ \text{TO}_{\text{Loc}} \\
\quad\quad ([_{\text{Thing}}\ \text{SOLDIER}], \\
\quad\quad\quad [_{\text{Position}}\ \text{AT}_{\text{Loc}} \\
\quad\quad\quad\quad ([_{\text{Thing}}\ \text{SOLDIER}], \\
\quad\quad\quad\quad\ [_{\text{Thing}}\ \text{BRIDGE}])])])]
\end{array}
$$

(ii)
```
(cause (act loc (soldier) (by march))
       (to loc (soldier)
           (at loc (soldier) (bridge))))
```

In the next sections we outline the aspectual properties of the LCS templates for verbs in the lexicon and illustrate how LCS templates compose at the sentential level, demonstrating how lexical aspect feature determination occurs via the same algorithm at both verbal and sentential levels.

# 4 Determining Aspect Features from the LCS Structures

The components of our LCS templates correlate strongly with aspectual category distinctions. An exhaustive listing of aspectual types and their corresponding LCS representations is given below. The `!!` notation is used as a wildcard which is filled in by the lexeme associated with the word defined in the lexical entry, thus producing a semantic constant.

(5) (i) **States:**

```
(be ident/perc/loc
    (thing 2) ... (by !! 26))
```

(ii) **Activities:**

```
(act loc/perc (thing 1) (by !! 26))
  or (act loc/perc (thing 1)
       (with instr ... (!!-er 20)))
  or (act loc/perc (thing 1)
       (on loc/perc (thing 2))
         (by !! 26))
  or (act loc/perc (thing 1)
       (on loc/perc (thing 2))
       (with instr ... (!!-er 20)))
```

(iii) **Accomplishments:**

```
(cause/let (thing 1)
  (go loc (thing 2)
     (toward/away_from ...))
          (by !! 26))
  or (cause/let (thing 1)
        (go/be ident
           (thing 2) ... (!!-ed 9)))
  or (cause/let (thing 1)
        (go loc (thing 2) ... (!! 6)))
  or (cause/let (thing 1)
        (go loc (thing 2) ... (!! 4)))
  or (cause/let (thing 1)
        (go exist (thing 2) ... (exist 9))
        (by !! 26))
```

(iv) **Achievements:**

```
(go loc (thing 2) (toward/away_from ...)
    (by !! 26))
  or (go loc (thing 2) ... (!! 6))
  or (go loc (thing 2) ... (!! 4))
  or (go exist (thing 2) ... (exist 9)
    (by !! 26))
  or (go ident (thing 2) ... (!!-ed 9))
```

The Lexical Semantic Templates (LSTs) of Levin and Rappaport-Hovav (To appear) and the decompositions of Dowty (1979) also capture aspectual distinctions, but are not articulated enough to capture other distinctions among verbs required by a large-scale application.

Since the verb classes (state, activity, etc.) are abstractions over feature combinations, we now discuss each feature in turn.

## 4.1 Dynamicity

The feature [+dynamic] encodes the distinction between events ([+dynamic]) and states ([∅dynamic]). Arguably "the most salient distinction" in an aspect taxonomy (Dahl, 1985, p. 28), in the LCS dynamicity is encoded at the topmost level. Events are characterized by `go`, `act`, `stay`, `cause`, or `let`, whereas States are characterized by `go-ext` or `be`, as illustrated in (6).

(6) (i) **Achievements:** *decay, rust, redden* (45.5)

```
(go ident (* thing 2)
   (toward ident (thing 2)
      (at ident (thing 2) (!!-ed 9))))
```

(ii) **Accomplishments:** *dangle, suspend* (9.2)

```
(cause (* thing 1)
   (be ident (* thing 2)
      (at ident (thing 2) (!!-ed 9))))
```

(iii) **States:** *contain, enclose* (47.8)

```
(be loc (* thing 2)
   (in loc (thing 2) (* thing 11))
   (by !! 26))
```

(iv) **Activities:** *amble, run, zigzag* (51.3.2)

```
(act loc (* thing 1) (by !! 26))
```

## 4.2 Durativity

The [+durative] feature denotes situations that take time (states, activities and accomplishments). Situations that may be punctiliar (achievements) are unspecified for durativity ((Olsen, To appear) following (Smith, 1991), inter alia). In the LCS, durativity may be identified by the presence of `act`, `be`, `go-ext`, `cause`, and `let` primitives, as in (7); these are lacking in the achievement template, shown in (8).

(7) (i) **States:** *adore, appreciate, trust* (31.2)

```
(be perc
   (* thing 2)
   (at perc (thing 2) (* thing 8)) (by !! 26))
```

(ii) **Activities:** *amble, run, zigzag* (51.3.2)

```
(act loc (* thing 1) (by !! 26))
```

(iii) **Accomplishments:** *destroy, obliterate* (44)

```
(cause (* thing 1)
       (go exist (* thing 2)
          (away_from exist (thing 2)
             (at exist (thing 2) (exist 9))))
       (by !! 26))
```

(8) **Achievements:** *crumple, fold, wrinkle* (45.2)

```
(go ident
   (* thing 2)
   (toward ident (thing 2)
      (at ident (thing 2)  (!!-ed 9))))
```

## 4.3 Telicity

Telic verbs denote a situation with an inherent end or goal. Atelic verbs lack an inherent end, though, as (1) shows, they may appear in telic sentences with other sentence constituents. In the LCS, [+telic] verbs contain a Path of a particular type or a constant (`!!`) in the right-most leaf-node argument. Some examples are shown below:

(9) (i) **leave**

```
(... (thing 2)
  (toward/away_from ...) (by !! 26))
```

(ii) **enter**

```
(... (thing 2) ... (!!-ed 9))
```

(iii) **pocket**

```
(... (thing 2) ... (!! 6))
```

(iv) **mine**

```
(... (thing 2) ... (!! 4))
```

(v) **create, destroy**

```
(... (thing 2) ... (exist 9) (by !! 26))
```

In the first case the special path component, `toward` or `away_from`, is the telicity indicator, in the next three, the (uninstantiated) constant in the rightmost leaf-node argument, and, in the last case, the special (instantiated) constant `exist`.

Telic verbs include:

(10) (i) **Accomplishments:** *mine, quarry* (10.9)

```
(cause
  (* thing 1)
  (go loc (* thing 2)
          ((* away_from 3) loc
            (thing 2)
            (at loc (thing 2) (!! 4)))))
```

(ii) **Achievements:** *abandon, desert, leave* (51.2)

```
(go loc
  (* thing 2)
  (away_from loc
```

```
        (thing 2)
        (at loc (thing 2) (* thing 4))))
```

Examples of atelic verbs are given in (11). The (a)telic representations are especially in keeping with the privative feature characterization Olsen (1994; To appear): telic verb classes are homogeneously represented: the LCS has a path of a particular type, i.e., a "reference object" at an end state. Atelic verbs, on the other hand, do not have homogeneous representations.

(11) (i) **Activities:** *appeal, matter* (31.4)

```
(act perc (* thing 1)
      (on perc (* thing 2)) (by !! 26))
```

(ii) **States:** *wear* (41.3.1)

```
(be loc (* !! 2)
      (on loc (!! 2) (* thing 11)))
```

# 5 Modifying the Lexicon

We have examined the LCS classes with respect to identifying aspectual categories and determined that minor changes to 101 of 191 LCS class structures (213/390 subclasses) are necessary, including substituting `act` for `go` in activities and removing Path constituents that need not be stated lexically. For example, the original database entry for class 51.3.2 is:

(12)

```
(go loc (* thing 2)
       ((* toward 5) loc
          (thing 2)
             (at loc (thing 2) (thing 6)))
       (by !! 26))
```

This is modified to yield the following new database entry:

(13)

```
(act loc (* thing 1) (by march 26))
```

The modified entry is created by changing `go` to `act` and removing the `((* toward 5) ...)` constituent.

Modification of the lexicon to conform to aspectual requirements took 3 person-weeks, requiring 1370 decision tasks at 4 minutes each: three passes through each of the 390 subclasses to compare the LCS structure with the templates for each feature (substantially complete) and one pass to change 200 LCS structures to conform with the templates. (Fewer than ten classes need to be changed for durativity or dynamicity, and approximately 200 of the 390 subclasses for telicity.) With the changes we can automatically assign aspect to some 9000 verbs in existing classes. Furthermore, since 6000 of the verbs were classified by automatic means, new verbs would receive aspectual assignments automatically as a result of the classification algorithm.

We are aware of no attempt in the literature to determine aspectual information on a similar scale, in part, we suspect, because of the difficulty of assigning features to verbs since they appear in sentences denoting situations of multiple aspectual types. Based on our experience handcoding small sets of verbs, we estimate generating aspectual features for 9000 entries would require 3.5 person-months (four minutes per entry), with 1 person-month for proofing and consistency checking, given unclassified verbs, organized, say, alphabetically.

# 6 Aspectual Feature Determination for Composed LCS's

Modifications described above reveal similarities between verbs that carry a lexical aspect feature as part of their lexical entry and sentences that have features as a result of LCS composition. Consequently, the algorithm that we developed for verifying aspectual conformance of the LCS database is also directly applicable to aspectual feature determination in LCSs that have been composed from verbs and other relevant sentence constituents. LCS composition is a fundamental operation in two applications for which the LCS serves as an interlingua: machine translation (Dorr et al., 1993) and foreign language tutoring (Dorr et al., 1995; Sams, 1993; Weinberg et al., 1995). Aspectual feature determination applies to the composed LCS by first assigning unspecified feature values—atelic [∅T], non-durative [∅R], and stative [∅D]—and then monotonically setting these to positive values according to the presence of certain constituents.

The formal specification of the aspectual feature determination algorithm is shown in Figure 1. The first step initializes all aspectual values to be unspecified. Next the top node is examined for membership in a set of telicity indicators (CAUSE, LET, GO); if there is a match, the LCS is assumed to be [+T]. In this case, the top node is further checked for membership in sets that indicate dynamicity [+D] and durativity [+R]. Then the top node is examined for membership in a set of atelicity indicators (ACT, BE, STAY); if there is a match, the LCS is further examined for inclusion of a telicizing component, i.e., TO, TOWARD, $\mathrm{FOR}_{\mathrm{Temp}}$. The LCS is assumed to be [∅T] unless one of these telicizing components is present. In either case, the top node is further checked for membership in sets that indicate dynamicity [+D] and durativity [+R]. Finally, the results of telicity, dynamicity, and durativity assignments are returned.

The advantage of using this same algorithm for determination of both verbal and sentential aspect is that it is possible to use the same mechanism to perform two independent tasks: (1) Determine inherent aspectual features associated with a lexical item; (2) Derive

Given an LCS representation L:

1. Initialize: T(L):=[∅T], D(L):=[∅R], R(L):=[∅D]
2. If Top node of L ∈ {CAUSE, LET, GO}
   Then T(L):=[+T]
      If Top node of L ∈ {CAUSE, LET}
         Then D(L):=[+D], R(L):=[+R]
      If Top node of L ∈ {GO}
         Then D(L):=[+D]
3. If Top node of L ∈ {ACT, BE, STAY}
   Then If Internal node of
   L ∈ {TO, TOWARD, $\mathrm{FOR_{Temp}}$}
         Then T(L):=[+T]
      If Top node of L ∈ {BE, STAY}
         Then R(L):=[+R]
      If Top node of L ∈ {ACT}
         Then set D(L):=[+D], R(L):=[+R]
4. Return T(L), D(L), R(L).

Figure 1: Algorithm for Aspectual Feature Determination

non-inherent aspectual features associated with combinations of lexical items.

Note, for example, that adding the path *to the bridge* to the [∅telic] verb entry in (3) establishes a [+telic] value for the sentence as a whole, an interpretation available by the same algorithm that identifies verbs as telic in the LCS lexicon:

(14) (i) [∅telic]: `(act loc (* thing 1) (by march 26))`

(ii) [+telic]:

```
(cause
   (act loc (soldier) (by march))
   (to loc (soldier)
      (at loc (soldier) (bridge))))
```

In our applications, access to both verbal and sentential lexical aspect features facilitates the task of lexical choice in machine translation and interpretation of students' answers in foreign language tutoring. For example, our machine translation system selects appropriate translations based on the matching of telicity values for the output sentence, whether or not the verbs in the language match in telicity. The English atelic manner verb *march* and the telic PP *across the field* from (1) is best translated into Spanish as the telic verb *cruzar* with the manner *marchando* as an adjunct:

(15) (i) E: The soldier marched across the field.
S: El soldado cruzó el campo marchando.

(ii)
```
(cause
   (act loc (soldier) (by march))
   (to loc (soldier)
        (across loc (soldier) (field))))
```

Similarly, in changing the *Weekend Verbs* (i.e., *December, holiday, summer, weekend,* etc.) template to telic, we make use of the measure phrase `(for temp ...)` which was previously available, though not employed, as a mechanism in our database. Thus, we now have a lexicalized example of 'doing something for a certain time' that has a representation corresponding to the canonical telic frame V *for an hour* phrase, as in *The soldier marched for an hour*:

(16)
```
(act loc (soldier) (by march)
         (for temp (*head*) (hour)))
```

This same telicizing constituent—which is compositionally derived in the *crawl* construction—is encoded directly in the lexical entry for a verb such as *December*:

(17)
```
(stay loc
   (* thing 2)
      ((* [at] 5) loc (thing 2) (thing 6))
   (for temp (*head*) (december 31)))
```

This lexical entry is composed with other arguments to produce the LCS for *John Decembered at the new cabin*:

(18)
```
(stay loc (john)
   (at loc (john) (cabin (new)))
       (for temp (*head*) (december)))
```

This same LCS would serve as the underlying representation for the equivalent Spanish sentence, which uses an atelic verb *estar* [4] in combination with a temporal adjunct *durante el mes de Diciembre*: *John estuvo en la cabaña nueva durante el mes de Diciembre* (literally, *John was in the new cabin during the month of December*).

The monotonic composition permitted by the LCS templates is slightly different than that permitted by the privative feature model of aspect (Olsen, 1994; Olsen, To appear). For example, in the LCS states may be composed into an achievement or accomplishment

[4] Since *estar* may be used with both telic *(estar alto)* and atelic *(estar contento)* readings, we analyze it as atelic to permit appropriate composition.

structure, because states are part of the substructure of these classes (cf. templates in (6)). They may not, however, appear as activities. The privative model in Table 2 allows states to become activities and accomplishments, by adding [+dynamic] and [+telic] features, but they may not become achievements, since removal of the [+durative] feature would be required. The nature of the alternations between states and events is a subject for future research.

## 7 Conclusion

The privative feature model, on which our LCS composition draws, allows us to represent verbal and sentential lexical aspect as monotonic composition of the same type, and to identify the contribution of both verbs and other elements. The lexical aspect of verbs and sentences may be therefore determined from the corresponding LCS representations, as in the examples provided from machine translation and foreign language tutoring applications. We are aware of no attempt in the literature to represent and access aspect on a similar scale, in part, we suspect, because of the difficulty of identifying the aspectual contribution of the verbs and sentences given the multiple aspectual types in which verbs appear.

An important corollary to this investigation is that it is possible to refine the lexicon, because variable meaning may, in many cases, be attributed to lexical aspect variation predictable by composition rules. In addition, factoring out the structural requirements of specific lexical items from the predictable variation that may be described by composition provides information on the aspectual effect of verbal modifiers and complements. We are therefore able to describe not only the lexical aspect at the sentential level, but also the set of aspectual variations available to a given verb type.

## References


Androutsopoulos, Ioannis. 1996. *A Principled Framework for Constructing Natural Language Interfaces to Temporal Databases.* Ph.D. thesis, University of Edinburgh.

Antonisse, Peggy. 1994. Processing Temporal and Locative Modifiers in a Licensing Model. Technical Report 2:1–38, Working Papers in Linguistics, University of Maryland.

Brinton, Laurel J. 1988. *The Development of English Aspectual Systems: Aspectualizers and Post-Verbal Particles.* Cambridge University Press, Cambridge.

Dahl, Östen. 1985. *Tense and Aspect Systems.* Basil Blackwell, Oxford.

Dorr, Bonnie J. 1997. Large-Scale Acquisition of LCS-Based Lexicons for Foreign Language Tutoring. In *Proceedings of the Fifth Conference on Applied Natural Language Processing (ANLP)*, Washington, DC.

Dorr, Bonnie J. To appear. Large-Scale Dictionary Construction for Foreign Language Tutoring and Interlingual Machine Translation. *Machine Translation*, 12(1).

Dorr, Bonnie J., James Hendler, Scott Blanksteen, and Barrie Migdalof. 1993. Use of Lexical Conceptual Structure for Intelligent Tutoring. Technical Report UMIACS TR 93-108, CS TR 3161, University of Maryland.

Dorr, Bonnie J., Jim Hendler, Scott Blanksteen, and Barrie Migdalof. 1995. Use of LCS and Discourse for Intelligent Tutoring: On Beyond Syntax. In Melissa Holland, Jonathan Kaplan, and Michelle Sams, editors, *Intelligent Language Tutors: Balancing Theory and Technology*. Lawrence Erlbaum Associates, Hillsdale, NJ, pages 289–309.

Dorr, Bonnie J. and Douglas Jones. 1996. Role of Word Sense Disambiguation in Lexical Acquisition: Predicting Semantics from Syntactic Cues. In *Proceedings of the International Conference on Computational Linguistics*, pages 322–333, Copenhagen, Denmark.

Dorr, Bonnie J., Dekang Lin, Jye-hoon Lee, and Sungki Suh. 1995. Efficient Parsing for Korean and English: A Parameterized Message Passing Approach. *Computational Linguistics*, 21(2):255–263.

Dorr, Bonnie J. and Mari Broman Olsen. 1996. Multilingual Generation: The Role of Telicity in Lexical Choice and Syntactic Realization. *Machine Translation*, 11(1–3):37–74.

Dowty, David. 1979. *Word Meaning in Montague Grammar*. Reidel, Dordrecht.

Dowty, David. 1986. The Effects of Aspectual Class on the Temporal Structure of Discourse: Semantics or Pragmatics? *Linguistics and Philosophy*, 9:37–61.

Grimshaw, Jane. 1993. Semantic Structure and Semantic Content in Lexical Representation. unpublished ms., Rutgers University, New Brunswick, NJ.

Hovav, Malka Rappaport and Beth Levin. 1995. The Elasticity of Verb Meaning. In *Processes in Argument Structure*, pages 1–13, Germany. SfS-Report-06-95, Seminar für Sprachwissenschaft, Eberhard-Karls-Universität Tübingen, Tübingen.

Jackendoff, Ray. 1983. *Semantics and Cognition*. The MIT Press, Cambridge, MA.

Jackendoff, Ray. 1990. *Semantic Structures*. The MIT Press, Cambridge, MA.

Klavans, Judith L. and M. Chodorow. 1992. Degrees of Stativity: The Lexical Representation of Verb Aspect. In *Proceedings of the 14th International Conference on Computational Linguistics*, Nantes, France.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation.* University of Chicago Press, Chicago, IL.

Levin, Beth and Malka Rappaport Hovav. To appear. Building Verb Meanings. In M. Butt and W. Gauder, editors, *The Projection of Arguments: Lexical and Syntactic Constraints.* CSLI.

Light, Marc. 1996. Morphological Cues for Lexical Semantics. In *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics.*

Moens, Marc and Mark Steedman. 1988. Temporal Ontology and Temporal Reference. *Computational Linguistics: Special Issue on Tense and Aspect*, 14(2):15–28.

Olsen, Mari Broman. 1994. The Semantics and Pragmatics of Lexical Aspect Features. In *Proceedings of the Formal Linguistic Society of Midamerica V*, pages 361–375, University of Illinois, Urbana-Champaign, May. In Studies in the Linguistic Sciences, Vol. 24.2, Fall 1994.

Olsen, Mari Broman. 1996. Telicity and English Verb Classes and Alternations: An Overview. Umiacs tr 96-15, cs tr 3607, University of Maryland, College Park, MD.

Olsen, Mari Broman. To appear. *The Semantics and Pragmatics of Lexical and Grammatical Aspect.* Garland, New York.

Passoneau, Rebecca. 1988. A Computational Model of the Semantics of Tense and Aspect. *Computational Linguistics: Special Issue on Tense and Aspect*, 14(2):44–60.

Pinker, Steven. 1989. *Learnability and Cognition: The Acquisition of Argument Structure.* The MIT Press, Cambridge, MA.

Resnik, Philip. 1996. Selectional Constraints: An Information-Theoretic Model and its Computational Realization. *Cognition*, 61:127–159.

Sams, Michelle. 1993. An Intelligent Foreign Language Tutor Incorporating Natural Language Processing. In *Proceedings of Conference on Intelligent Computer-Aided Training and Virtual Environment Technology*, NASA: Houston, TX.

Sams, Michelle. 1995. Advanced Technologies for Language Learning: The BRIDGE Project Within the ARI Language Tutor Program. In Melissa Holland, Jonathan Kaplan, and Michelle Sams, editors, *Intelligent Language Tutors: Theory Shaping Technology*. Lawrence Erlbaum Associates, Hillsdale, NJ, pages 7–21.

Siegel, Eric V. and Kathleen R. McKeown. 1996. Gathering Statistics to Aspectually Classify Sentences with a Genetic Algorithm. Unpublished MS (cmp-lg/9610002)., Columbia University, New York, NY.

Slobin, Dan I. and Aura Bocaz. 1988. Learning to Talk About Movement Through Time and Space: The Development of Narrative Abilities in Spanish and English. *Lenguas Modernas*, 15:5–24.

Smith, Carlota. 1991. *The Parameter of Aspect*. Kluwer, Dordrecht.

Tenny, Carol. 1994. *Aspectual Roles and the Syntax-Semantics Interface*. Kluwer, Dordrecht.

Verkuyl, Henk. 1993. *A Theory of Aspectuality: The Interaction Between Temporal and Atemporal Structure*. Cambridge University Press, Cambridge and New York.

Weinberg, Amy, Joseph Garman, Jeffery Martin, and Paola Merlo. 1995. Principle-Based Parser for Foreign Language Training in German and Arabic. In Melissa Holland, Jonathan Kaplan, and Michelle Sams, editors, *Intelligent Language Tutors: Theory Shaping Technology*. Lawrence Erlbaum Associates, Hillsdale, NJ, pages 23–44.