Granularity in the Cross-linguistic Encoding of Motion and Location

Miriam van Staden
University of Amsterdam & MPI, Nijmegen

Bhuvana Narasimhan
MPI, Nijmegen

Abstract
In this study, we explore three ways in which the notion of "granularity" emerges from the study of cross-linguistic event semantics. The first interpretation of granularity has to do with event segmentation for linguistic expressions. Where humans place event boundaries varies depending on the language and cultural setting in which the event is encoded. Second, within the set boundaries of a 'single event' in time there are many elements that could all receive expression, and again languages show variation, both with respect to which elements are mentioned at all, our second interpretation of granularity, and the specificity with which these elements are characterized, our third interpretation. It remains a matter for empirical research to discover whether these differences in linguistic expression reflect or inform different representations of event boundaries and event classification for non-linguistic purposes as well.

1. Introduction
In this paper we look at similarities and differences in how people linguistically encode events of motion and location. In particular, we examine habitual, colloquial descriptions of caused motion into containment (as in sentences such as he put the book into the bag), to explore how languages differ with respect to the segmentation and classification of events. We suggest that, while the ability to segment the continuum of experience and perception into event units and talk about them in more or less fine-grained ways is universal, there are differences between speakers of different languages in the level of granularity at which events are typically referred to in linguistic descriptions. Three interpretations of granularity appear particularly relevant. First, there are cross-linguistic differences
with regard to where event boundaries are placed in linguistic descriptions. Second, within the boundaries of an event, there are systematic differences in the elements that are given expression. Finally, there is considerable cross-linguistic variation in the level of detail that languages care about when they do express particular elements of the event. The last two notions of granularity reflect more or less fine-grained ways in event classification; the first shows a coarser or finer grain in the partonomic level of event description. We begin by characterizing in further detail, each of the notions we have just introduced, relying heavily on the excellent overview of theorizing and empirical research on event structure provided in (Zacks & Tversky, 2001a).

Any description of a motion event can be characterized in terms of a core set of components including **figure, motion, path, ground, manner, cause** (Talmy, 1985, 1991). Hence, in sentence (1) the noun phrase *the book* is the figure and *the bag* is the ground, the preposition *into* describes the path, and the verb *slide* encodes cause and manner of motion.

(1) **he slid the book into the bag**

In addition, in caused motion events, there is a **causer** (*he* in the sentence above). Languages encode these components of the event in a variety of ways depending on the lexical and constructional resources in the language, and Talmy (1985, 1991) suggests that languages differ systematically in how they incorporate components such as manner in the encoding of motion events. He observes that some languages typically encode the path information in the verb, e.g. Spanish, while other languages like English typically encode manner information in the verb. It also appears that languages differ in which components of the event are typically encoded. Thus, Spanish encodes manner of motion strikingly less often than English (**check Slobin, 2000**).

In theorizing about how events might be perceived and conceptualized, Zacks & Tversky (2001) suggest that components such as figure, motion, path, ground in linguistic descriptions point to an underlying structured representation of events on which people rely in talking about events (pp.10-11). The basic building block of (motion) events ‘should be temporal units in which the figure, motion, path, and ground are constant’ and a change in the motion, path, or ground relative to the figure would mark the boundary where a new (atomic) event begins (pp.9-10). Thus, a general motion event such as ‘going skiing’ can be partitioned into segments such as ‘riding the ski lift’, ‘getting off the lift’ and ‘continuing skiing’, ‘turning at the base of the ski jump’, and so on (p.10). A change in the figure,
however, typically starts a new series of atomic events that together form an ‘intentional action’. A
series of intentional actions together yield a ‘script’. In this manner, the smaller event units can be
grouped into larger units to form a partonomic hierarchy. For instance, the activity of ‘going skiing’
might itself constitute a subpart of an event such as ‘taking a winter sports course’ which might then be
part of an event at a broader time scale such as ‘becoming a ski instructor’. At the other end of the
hierarchy, the event of ‘getting off the lift’ can have further subparts such as ‘stepping off the lift’
(Barker & Wright, 1954). In this partonomic hierarchy Zacks & Tversky furthermore identify a
‘privileged partonomic level’ which includes behavior episodes such as ‘a boy going home from
school’, or ‘a girl exchanging remarks with her mother’ (cf. Barker & Wright, 1954), or, in another
approach, scenes in a script (e.g. scenes in a ‘restaurant script’ include ‘entering,’ ‘ordering,’ ‘eating,’
etc.) (cf. Schank & Abelson, 1977). When presented with actions at a subordinate level, people make
inferences up to the scene level, but are unlikely to make downward inferences to the subordinate level
when presented with information at the scene level (Abbott, Black, & Smith, 1985). Zacks & Tversky
suggest that at such a level in the partonomic hierarchy ‘cognition is particularly fluent’ (p.10).

Events can also be characterized by a taxonomic hierarchy that is based on kind-of rather than
part-of relationships. Thus, frisbee golf is a kind of golf, which in turn is a kind of sport (Zacks &
Tversky, 2001a:5). Some evidence for a preferred basic level on a taxonomic hierarchy exists as well.
For instance, Morris & Murphy (1990) found that participants responded fastest to basic-level labels
when given an excerpt from event descriptions (e.g. ‘scream during the scary parts’) and asked to
verify a category label at the subordinate level (‘horror movie’), basic level (‘movie’), or superordinate
level (‘entertainment’). Similarly, ‘going skiing’ could be subclassified further as ‘going downhill
skiing’ vs. ‘going cross-country skiing,’ and ‘going downhill skiing’ might be differentiated further
into ‘bunny-slope skiing’ versus ‘mountain-slope skiing.’

Within any cultural-linguistic setting, people can talk about events at different temporal levels
and different degrees of specificity. Which level of specificity is chosen depends to a large extent on
the particular setting in which the event is described. When asked a question such as what did you do
today, it is likely that people give an answer that is at a higher temporal resolution (I went to the
theatre) than that of an answer to a question such as what did you do last year? (I took at trip to
Guatemala). But if the conversation takes place during a dinner party the answer will also have a
higher temporal resolution then if it takes place during a trans-Atlantic collect-call. Thus depending on
the context, interlocutors make judgments on the appropriate level of event description. This choice for a particular temporal resolution is part of a set of more general maxims governing discourse, which relate to the expected truthfulness, informativeness, and relevance of utterances in verbal interaction (Grice 1975). For instance, if you are looking for somebody in a large building and you ask someone where is Sally?, you expect the answer to be as precise as is necessary for you to find Sally, but no more precise than this. An answer like she is in the building when the speaker actually knows she is in the library on the eleventh floor is too poor, while an answer like she is in the newly renovated library on the eleventh floor near the window at the second desk from the left sitting in a red chair, reading a book on conversational implicatures may be unnecessary prolific. Again, what we judge to be adequate depends very much on the situation in which the utterance is made, for instance, if the library is large and many of the desks are hidden from view it may actually be very helpful to know that Sally is near the window.

Yet what is considered the correct level of informativeness for a given situation also depends on the cultural and linguistic setting in which the discourse occurs. For instance, in motion and locative descriptions in Tidore speakers will almost invariably indicate the direction in which an entity is moving or is located, even when to the English ear this may appear entirely redundant. If in a small room there is only one table and someone asks where her mug is, the answer is likely to be something along the following lines:

(2) ngona na-mok=ge katina toma meja ntina=ge
    you your-mug=there be.landwards LOC table landward.one=there
    ‘your mug is in a landward location on the landward table’

Not just is the information that the hearer must turn ‘landward’ to find the mug mentioned twice, it also appears rather redundant given the fact that there is only one table, and this table is clearly visible to both speaker and hearer. For an English speaker, this may sound like unnecessarily prolific, but for a speaker of Tidore anything less is too imprecise. While the sentence without the directional and locational may be grammatical, speakers will observe that ‘we just don’t say it that way’, or that ‘it is not really clear where the mug is now.’ Similarly, as we demonstrate below, Tzeltal motion and locative descriptions contain much detail on the precise orientation and shape, size, etc. of the figure or ground object involved. The English speaker may wonder whether it is at all relevant in non-contrastive situations to specify that the ball is not just in a box, but in a box with a narrow opening or in an oblong
box. And if the ball is on a mat, does it matter that this mat is made of flexible material? Yet, Tzeltal apparently does care for these distinctions as evidenced by the large number of INSERT verbs that convey this type of information. Even if a more generic description is at all possible, speakers consistently choose the more specific one.

Thus, in addition to a privileged level of granularity in the sense used by Zacks & Tversky, we can also identify a ‘basic level’ when we refer to what is typically encoded in descriptions of comparable everyday situations like those above, where we assume that the informational needs are similar (‘where is Sally?’, ‘where is my cup?’). The basic maxims governing verbal interactions are also the same, but the ideas about just what this means are different. These differences result in different communicative strategies, including, as we shall argue systematic differences across languages in the granularity of description at the ‘basic level’ in all three senses of the term that we describe in this paper: what constitutes an event; which elements in the event deserve mention; and with what richness of detail are these elements expressed?

Summarizing, there is evidence from the literature that events can be characterized in terms of hierarchies, either partonomic (involving partitioning events into constituent elements, as in Talmy’s work, or into temporally arranged parts as described by Zacks & Tversky) or taxonomic (classifying events into larger or smaller subtypes). Further, in describing events, people are likely to zoom in at a particular grain level, depending on the context. In the remainder of this paper, we explore cross-linguistic data that suggest that the level on the hierarchy (either partonomic or taxonomic) at which a speaker chooses to describe an event also varies, within semantic domains, according to the specific language in which the event is encoded. We present data from a number of different languages: English, Dutch, Hindi (Indo-European, spoken in N. India), Tidore (Papuan, spoken in Eastern Indonesia), Tzeltal (Maya, spoken in Mexico), Kalam (Papuan, spoken in mainland PNG) and Kilivila (Austronesian, spoken in insular PNG) with respect to caused motion expressions such as he put the ball into the box.

2. The unit of analysis

Our unit of analysis is the predicate, which we define as a complex of one or more relational entities in the clause with a predicative function (including simple verbs, verb + particle constructions, verb + light verb constructions, etc. but excluding modifying elements such as adverbials, or the secondary
predicate in depictive constructions). Thus, the verb *hobble* in *she hobbled slowly* constitutes a predicate, whereas *hobble out* is the predicate in *she hobbled out*. Note that we do not consider the direct object of the verb (*the box* in the phrase *kicked the box*) to be part of what we define as the predicate. Further, we focus in our discussion on those predicative constructions that occur in ‘basic level’ descriptions, defined earlier as unmarked, habitual descriptions of situations that are frequent in natural discourse. In the context of caused motion into containment these would be the typical answers to questions such as: ‘what happened to figure X?’, ‘where is X?’ or ‘what does causer A do to figure X?’ For these expressions we can then ask: which part of the continuum of experience and perception is selected for expression, and how abstractly or richly is it described?

### 3 Partitioning events in language

The first interpretation of granularity hinges on the question of where humans place event boundaries in their linguistic descriptions. We have some evidence from psychology that within cultural groups (predominantly northern Americans) there is strong agreement on what constitute natural event boundaries (Newton, 1973; Zacks & Tversky, 2001a). Further, as discussed in the introduction, sequences of smaller or more elementary events can build into larger structures to form a hierarchical organization (Byrne, 2002; Zacks et al., 2001; Zacks & Tversky, 2001b). Three levels of granularity are distinguished (Zacks & Tversky, 2001a). At the lowest level, events are perceived physical changes in the environment. These are the smallest segments in the continuum of perception and experience that are conceived of as single units. Intentional actions form the intermediate level and concern goal directed actions and causal relations between physical changes. These intentional actions can be grouped to give scripts, the highest level. Scripts are recursive in the sense that they can be parts of other scripts, and this distinguishes them from intentional actions. Since the perception of goal directed actions and causal relations are universals of human cognition and central in our conception of changes in the environment, it may be expected that all languages will have expressions of this type of event.

It is not necessarily the case, however, that humans in all cultures will group the same set of physical changes as one intentional action in describing them. Where a unitary event starts and what constitutes the end-point may be different for different languages and cultures. Von Stutterheim et al. (2002) showed that speakers of German are more likely to express an endpoint of a motion than speakers of English or Spanish. Where the latter would describe a scene as ‘the boat was sinking’, the
Germans would say ‘the boat sank to the bottom of the ocean’, even when the endpoint of the event was not visible. The explanation for this difference is in preferences is given as the absence of a productive progressive aspect in German that allows for a focus of the event as ongoing, which both English and Spanish do have. In a similar way, speakers of Tidore add inceptions of events. When shown a video-clip of a man chopping wood, they are likely to describe this as follows:

(3) $Nau=ge\; oro\; peda\; tola\; luto$
    $\text{boy=there}\;\text{fetch machete}\;\text{chop fire.wood}$
    ‘The boy fetches a machete (and) chops fire wood’.

Note that this is regardless of whether the actual picking up of the machete is shown. English speakers clearly do not regularly do this. Pawley (1987:346) shows that for Kalam, a Papuan language of Melanesia, intentional action are systematically reported as: 1. movement to scene of first action; 2. action; 3. movement from scene of 2 to present or final scene; 4. action(s) at present or final scene. Hence, an event, which in English would be encoded as *I gather firewood* would, in Kalam, be expressed as ‘I go (1.) wood strike (2.) get come (3) put (4)’. This type of event report is in fact very common in the Papuan languages of Melanesia, as well as some Austronesian ones that perhaps adopted this strategy through language contact (cf. van Staden & Senft, 2001), and is possibly related to the general avoidance of having more than one full noun phrase or more than two overt (pronominal) arguments per predicate-argument structure so that all ditransitive actions and all actions involving manipulation of multiple objects are distributed over more than one predicate (de Vries, 2003; Du Bois, 1985, 1987; Heeschen, 1998). In this interpretation of granularity, languages can be shown to be different in where they place the boundaries for event report.

In events of caused motion into containment we find similar language-specific differences in how events are partitioned. An English speaker will encode in a verb+preposition/particle predicate, the causer manipulating the object, the path of the motion and the result state in which the figure is contained by the ground: *he put the ball into the box*. A speaker of Kilivila (Austronesian, spoken in insular Papua New Guinea) will first express the event where the causer takes up the object and then goes on to describe the path of the object, the causal relation and the topological relation, followed possibly by the end state, all in a single clause (Senft, pc):
These are typical descriptions of caused locatives in natural discourse. The prosodic contour shows them to be single units, and indeed in repair the entire sequence will be repeated and never just part of it, so that they function in every respect as single clauses. When a single verb clause is deemed grammatical at all, native speakers of Kilivila will consider it ‘foreigner talk’ (Senft, p.c.).

Tidore (Papuan, spoken in Eastern Indonesia) similarly has serial verb constructions that express ‘causer picks up figure’ and ‘figure is placed inside ground’. Consider the following descriptions of caused change of location into containment events:

(6) **Una oro fanai kam gure toma oti ngge ma-doya**

he fetch bait ‘contents’ put LOC perahu there its-inside

‘He fetched the bait put (poured) them inside (into) the perahu’

(7) **Ngona musti no-oro goroho ngge gure toma tempayang nde**

you must you-fetch oil there put LOC container here

   ma-doya koliho
   its-inside back

‘You must fetch the oil and put it back inside this container’

Equivalent descriptions of similar scenes in a language such as Hindi (Indo-European, spoken in Northern India) cannot express a similar partitioning of events within the main clause. Consider the following equivalents of the Tidore examples in (6) and (7) above. In Hindi such complex events (bringing+putting or carrying+bringing) only be encoded using a participial verb phrase in conjunction with the main clause:

(8) **tel-0 laakar botal mE Daal do**

oil-NOM bring-CONJ bottle-LOC put give-IMP

‘Having brought the oil, put (it) in the bottle.’

(9) **us-ko uThaakar laao.**

it-ACC lift/carry-CONJ bring-IMP

‘Lift/carry it and bring (it).’
The availability of serial-verb constructions as part of the grammatical toolkit of the language can only be part of the reason why there are systematic differences in event report between Kilivila and Tidore on the one hand, and Hindi on the other. First of all, it is not the case that these descriptions are always in the form of a serial verb construction. For example, the following description from Tidore shows the FETCH event in a separate clause from the INSERT event, separated by conjunction *la* ‘so (that)’:

(10) Oro una toma Cobo gosa ino la gure una toma kurunga
    fetch he LOC C. carry this.way so put he LOC cage
    ma-doya ma
    its-inside just
    ‘(They) fetched him from Cobo carried here so that they just put him in a cage’

Moreover, Hindi does have a verb+verb structural template, but the semantics associated with it does not allow the insertion of full verbs like in Tidore. In Hindi V+V templates (e.g. *Daal de* ‘put give’) are available and used with high frequency only to express aspectual distinctions (e.g. completion, inception) related to either one of the two events expressed in Tidore or Kilivila. The second verb ‘give’ is not a fully lexical verb; the lexical verbs for ‘put’ and ‘lift’ cannot be combined:

(11) gend-ko Dibbe-mE Daal do
    ball-ACC box-LOC put give-IMP
    ‘Put the ball in the box.’

(12) gend-ko uThaaao do
    ball-ACC lift give-IMP
    ‘Pick up the ball.’

(13) *gend-ko Dibbe-mE uThaa Daalo
    *ball-ACC box-LOC lift put-IMP
    *‘Lift the ball put into the box.’

So even if we disregard the question of construction type or number of predicates per clause there are clearly differences between Tidore and Hindi. Depending on how one interprets these differences Tidore event reports are more coarse grained than Hindi, in the sense that they systematically select larger portions out of the continuum of experience and perception as single event units. At the same time they are more fine grained to the extent that they give more detail on the internal structure of the
event, if it can be assumed that the Hindi description metonymically refers to more than is lexicalized. The description of the internal structure is the focus of the following two sections.

In summary, we suggest that the ability to partition events for the purpose of talking about them is a cognitive ability that all humans share, and that when pushed speakers will be able to play with these event boundaries and verbalize events at a coarser or finer grain level. But the basic level of granularity that speakers typically use is not fixed across languages. And we find that the grammatical and lexical resources of the language to some extent reflect the default level of granularity, for instance in the existence of serial verb constructions that allow for both events to be expressed in a single clause. While this suggests a structuring of events for the purposes of speaking (cf. Slobin, 1985, 1991), whether the linguistic encoding of events influences the partitioning of events for non-linguistic purposes is a matter for further research.

4 Levels of generality in event classification

Once the boundaries of an event are determined, the events can be classified taxonomically on the basis of the number of elements that are encoded by the predicate at the clausal level, and the refinement in the expression of these elements. In this section we discuss each in turn. Consider again the event of caused change of location into containment. It has been shown, that for such types of events, languages lexicalize similar components or elements in motion event descriptions (e.g. sentence the boy rolled the ball into the box ) (Talmy, 1985, 1991):

- Figure: ball
- Ground (Source/Goal/Midpoints): the box
- Path: into
- (Caused) Motion: roll
- Manner: roll

Thus, the expression of a caused motion into containment potentially encodes at least a motion or translocation, a manner in which the motion occurs, a directionality of the motion, a figure object that is inserted, a ground object that is the container, and a causer. We refer to these as the elements in the event, to avoid confusion with the parts of the event that are related through connection in time.\{1\}
These elements in motion descriptions are not always all expressed. For instance in *the boy left the house* manner is not expressed, and in *she ran out* there is no *ground* expression. Again languages are shown to be different in the resources they have to express elements of the motion description, in particular in the predicate, as well as in how typically make use of these resources to express the various elements in a motion description. A description can be said to be more fine-grained if it expresses more elements in the predicate. More fine-grained descriptions show a more precise taxonomic classification of events. We mentioned earlier how the game *frisbee golf* is characterized as a subtype of *golf* based on the specification of one of the elements of golf, viz. the type of object it is played with (cf. Zacks & Tversky, 2001). Similarly, *run and walk* are more specific than *move* because they express aspects of the manner of motion, and *descend or move up* are more specific because they express the directionality of the motion. The English verb for caused motion into containment *put* is thus highly general it expresses aspects neither of figure or ground, nor of the kind of topological relation that is brought about. In English this is expressed in the prepositional phrase introduced by a basic preposition or by a particle or relational noun:

(14) *He inserted the books into the bag.*
(15) *He put the books inside the bag.*
(16) *He crammed the books in.*

Hindi, too, uses a single verb in conjunction with a ground denoting phrase. Two different construction types are found, one with and one in which a spatial nominal forms a possessive construction with the ground object (‘box’s inside’) as in (17), and one in which a locative case enclitic marks the containment relation directly on the ground object (‘box-in’), as in (18):

(17) *us-ne is-ko Dibbe-ke andar ghusaayaa.*  
he-ERG it-ACC box-GEN inside enter-make  
‘He inserted it inside the box.’

(18) *us-ne is-ko Dibbe-mE ghusaayaa.*  
he-ERG it-ACC box-LOC(in) enter-make  
‘He inserted it in the box.’

Dutch is an interesting language in this respect since it has first of all a choice of predicates depending on depending on the classification of the figures as canonically ‘sitting’, ‘standing’ or
‘lying’. In addition, Dutch also has a generic verb *stoppen* ‘put, insert’ that can be used only for containment relations but then is impartial to the kind of figure that is located:

(19)  
\[ \text{Hij legt/stopt de bal \textit{in} de doos} \]  
he lies/puts the ball in the box  
‘He puts the ball in the box.’

(20)  
\[ \text{Hij legt/*stopt de bal \textit{op} tafel} \]  
he lies/puts the ball on table  
‘He puts the ball (lying) on the table.’

(21)  
\[ \text{Hij zet/*legt het kopje \textit{in} de kast} \]  
he stands/lies the cup in the cupboard  
‘He puts the cup standing in the cupboard.’

Tzeltal and Tidore show a refinement in the predicate not often taken into account in motion descriptions. In addition to a path element, they also express the directionality or orientation of the motion with respect to a deictic center. In neither language this element is obligatory, but the fact that in Tidore the directional verbs are among the ten most frequent verbs in the language, along with verbs for ‘say’, ‘put’, ‘move/go’ and ‘make’, shows how frequent they are in use.\(^2\) The constructions used in Tidore and Tzeltal are like the Hindi constructions described above. They use relational nouns in possessive constructions in combination with a general (locative) preposition as in examples (22) and (24); or the locative (or Tzeltal general preposition) is used on its own (examples (23) (25)). However in both types of constructions, it is quite natural and common to further specify the direction of the motion, for example ‘seawards’ or ‘landwards’ in Tidore and ‘awaywards’ or ‘coming’ in Tzeltal:\(^3\)

(22)  
\[ \text{Una wo-gure \textit{ena} hoo \textit{toma gardus ma-doya}} \]  
he he-put it \textit{seawards} LOC box its-inside  
‘He put it seawards in the box’

(23)  
\[ \text{Dadi rofu \textit{ena} gure \textit{isa} toma hono} \]  
so weed it put \textit{landwards} LOC bowl  
‘So the weed it put landwards in a bowl’

(24)  
\[ \text{ya k-otzes \textit{bel} to \textit{karton}} \]  
INC 3E insert 3A \textit{awaywards} AT inside of box  
‘I insert awaywards at the inside of the box’
Thus, while deictic notions encoded in *go* and *come*, or vertical direction in *ascend* and *descend* can also be stacked in expressions in Hindi ("wo hamaarii taraf andar aayaa ‘he came inside, towards us’) and English (e.g. *he put the book awaywards on the shelf* or *she inserted the pencil in the hole in the upward direction*), these are neither natural nor habitual expressions of caused motion into containment in either language. Tzeltal and Tidore do encode simple motion to a goal resulting in a containment relation (e.g. *insert in the cup*) just as in Hindi and English. However, owing to the possibility of incorporating a directional verb with a main verb in a single construction, it also typically makes further distinctions within the class of motion-into-containment events.

In examining the finer subdivision of motion events on the basis of distinctions typically made in the encoding of directionality of motion in Tidore and Tzeltal, we focused on the number and kinds of elements that are habitually encoded in the predicate in motion event descriptions. However, even when two languages both express (aspects of) a particular element in the predicate, this does not mean that they are equally precise in the event description. Tzeltal, for example, chooses from two directional verbs, while Tidore has six. Apart from a centripetal and a centrifugal verb, it has directional verbs for ‘upward’, ‘downward’, ‘landwards’ and ‘seawards’. Tzeltal, however, is as we illustrate below, much more specific in terms of the specification of goal argument properties by the verbs, while the verb in the caused motion description in Tidore is highly unspecific. In fact, there is not a single dedicated verb that means ‘to cause a relation in which one object is in or inside another object’ (i.e. a word that could be glossed as *insert*). Only a few caused locative verbs give semantic detail with respect to the kind of figure or the kind of ground that is involved. For instance, *sose* is a causative locative verb used only for spreading out tablecloths on tabletops. Apart from these, Tidore only has *gure* ‘put’ and *ten* ‘put, place’, and a derived causative verb *somasusu* ‘cause to be entered’.

Depending on how much information it encodes about the characteristics of the goal (in the form of selectional restrictions), a verb can act as a zoom lens classifies events based on (figure and/or ground) object properties in more or less fine-grained ways. This is then our third and final interpretation of ‘granularity’ in motion descriptions. Predicates in different languages have interestingly different characteristics in this respect. In a language such as Hindi, the mono-morphemic verbs of caused motion into containment include *bharnaa* ‘fill (liquid/aggregates)’, *ghusaanaa* ‘insert,
fill (non-liquid masses) stuff’, ghuseD\n\n\nghus\n\nh\naa
\n‘cram’, and ThUUs\naa
\n‘force down, cram in’. While the latter three verbs imply force-dynamic interactions between the participants involved in the action, there is no semantic specification of the spatial characteristics of the goal, other than that it is a (3D) container. In English, caused motion into a container can be expressed by mono-morphemic verbs such as cram, fill, stuff, insert, dip, dunk, pierce, etc. as well as a range of verbs derived from nouns such as bag, bin, bottle, box, can, tin, crate, garage, house, jail, kennel, pocket etc. (Levin, 1993). The former set of verbs imply something about the manner and potential end result of the action, however, they are not very informative about the properties of their objects, with the exception of dip and dunk which imply that the goal is a liquid. In contrast, the set of denominal verbs provide highly specific information about the typical shape, size, and even material (e.g. bottles are usually made of glass, tins of metal) of the entities which might function as the goal of the caused motion event.

Now, let us consider the goal specifications of a range of INSERT verbs in Tzeltal (Brown, 1994). While a verb of caused motion into containment such as och-es places little restriction on the goal (other than that it has an ‘inside region’, 1994:769), other verbs are more particular. For instance, tik’ requires a goal which is a 3D bounded space with a narrow opening (e.g. bowl, narrow-mouthed gourd, cardboard tube), lap suggests a flexible goal (cloth, mat), and t’um-an implies liquid in a container (Brown, 1994: 769). Other predicates imply various properties of the container goals: chejp-an, a bag-like container (e.g. netbag), pajch-’an, a wide-mouthed container (e.g. bowl), and wajx-an, a tall oblong-shaped container (e.g. bottle) (Brown, 1994: 760-769).

Going from global to local, we can classify verbs of caused motion into containment into verbs which:

- specify only caused motion, with containment specified by a relational noun such as INSIDE or left to pragmatic inference (e.g. when a general locative is used, e.g. ‘put LOC bag’ as in Tidore),
- specify that the goal is a container,
- imply characteristics of the container including shape, width of the opening, rigidity, physical state (solid vs. liquid), and
- verbs which name a class of containers.
The lexicalization patterns and the constructional resources available in a language thus interact to produce, at the level of the predicate (complex): (a) very fine-grained motion descriptions in Tzeltal (e.g. insert-awaywards in a narrow-mouthed container, insert-upwards in a tall, oblong-shaped container); (b) descriptions of intermediate specificity with denominal verbs in English (e.g. bottle the pickles in the jar, box the books in cardboard cartons) or Tidore (e.g. ‘put seawards LOC the jar’s inside’), (c) coarse-grained descriptions in Hindi and English (e.g. stuff inside the bag, dunk into the coffee), and sparse descriptions where containment is left to pragmatic inference entirely as in Tidore (‘put the book LOC the bag’).{5}

In this section, we have described the subclassification of motion event descriptions in terms of distinctions made on the basis of features such as directionality and the properties of the ground object. Languages pack information into verbs and clauses to different degrees, and there is both cross-linguistic and intra-linguistic variation in this respect. It remains to be seen how we can characterize the scope and limits of this variation in a systematic way.

5. Conclusions

In this chapter we have shown that there is considerable variation in terms of where event boundaries are placed and how richly the event is characterized in terms of its constituent elements. Much further research is required to determine whether there is a small number of granularity levels in the way languages encode information lexically and combine them in specific construction types, or whether there is continuous variation in this respect. Thus, while taxonomic and partonomic hierarchies might underlie the representation of events for speakers of all languages, a number of factors underlie the selection of the particular levels which speakers select for the segmenting and categorizing of the continuum of experience and perception. We suggest that one of these factors is the particular preferences that speakers of different languages have for encoding events at a particular granularity for unmarked, basic-level descriptions of the event. Such preferences may vary intra-linguistically as well. Further cross-linguistic research is required to investigate the issues we have raised, as well as some interesting implications of this variation, including the extent to which language-specific preferences might impinge upon non-linguistic cognition and vice-versa.
**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>CONJ</td>
<td>conjunctive</td>
</tr>
<tr>
<td>IMP</td>
<td>imperfective</td>
</tr>
<tr>
<td>ERG</td>
<td>ergative</td>
</tr>
</tbody>
</table>


*Pragmatics, 1*, 7-26.


NOTES

* Many thanks to Penny Brown and Gunter Senft for their generosity in sharing their knowledge of respectively Tzeltal and Kilivila and providing us with many examples. Also, we are very grateful to the members of the Acquisition and Language & Cognition groups at the Max Planck Institute in Nijmegen for their input on the issues discussed in this paper. The views expressed here are our own, as well as any errors.

1 Note that the elements in event report that are distinguished in Talmy’s approach to event partonomies apply to each level in the hierarchy. Script level expressions, intentional actions and physical changes can all express motion events, and for all figures, grounds, manners, etc. can be identified.

2 Directional verbs in Tidore implicate but do not entail motion. ‘Fact of motion’ may be expressed separately by the verb tagi ‘move, go’, but this element, too, is not obligatory in a motion event.

3 Tzeltal also encodes paths such as ENTER, EXIT, COME, GO, ASCEND, and DESCEND in bound verbal morphemes (directionals). These directionals can occur together with the main path verb in the same construction (e.g. ‘put-enter’), thus allowing paths to be (optionally) stacked in a single clause.

4 Tzeltal locative descriptions may also specify properties of the figure. (Brown, 1994; see also Talmy, 1985 for related observations with respect to Atsugewi)

5 Note that we are talking only about information expressed by predicates at the constructional level (verbs, particles, directionals and their combinations); if we include information encoded in the noun phrases (e.g. the bag, the cupboard, etc.), then English and Hindi also specify detailed information about the properties of the goal.
Summary (= abstract)

In this study, we explore three ways in which the notion of "granularity" emerges from the study of cross-linguistic event semantics. The first interpretation of granularity has to do with event segmentation for linguistic expressions. Where humans place event boundaries varies depending on the language and cultural setting in which the event is encoded. Second, within the set boundaries of a ‘single event’ in time there are many elements that could all receive expression, and again languages show variation, both with respect to which elements are mentioned at all, our second interpretation of granularity, and the specificity with which these elements are characterized, our third interpretation. It remains a matter for empirical research to discover whether these differences in linguistic expression reflect or inform different representations of event boundaries and event classification for non-linguistic purposes as well.

Two-line biographies

Bhuvana Narasimhan is a linguist specialised in first language acquisition of verb argument structure, in particular in Hindi and Tamil.

Miriam van Staden is a linguist working on serial verb constructions and event report in Papuan and Austronesian languages.

Word count

Plain text: 5,767;
Including abstract, notes and references: 6,785