P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18

# 3 The macro-event property

## The segmentation of causal chains

*Jürgen Bohnemeyer, N. J. Enfield, James Essegbey, and Sotaro Kita*

### 1 Towards a semantic typology of event segmentation

Semantic typology is the study of linguistic categorization. In the simplest case, semantic typology investigates how an identical perceptual stimulus is categorized across languages. The problem examined in this article is that of event segmentation. To the extent that events are perceivable,[1] this may be understood as the representation of dynamic stimuli in chunks of linguistic code with categorical properties. For illustration, consider an example from a classic study on event cognition (Jenkins, Wald and Pittenger 1986): a woman prepares a cup of tea. She unwraps a tea bag, puts it into the cup, gets a kettle of water from the kitchen, pours the water into the cup, etc. This action sequence can be diagrammed schematically as in fig. 3.1.

It is conceivable that at some level of "raw" perception – before the onset of any kind of categorization – the action sequence is represented as a continuous flux. But it is hard to imagine how higher cognitive operations of recognition and inference could operate without segmenting the stream of perceived activity into units that are treated as instances of conceptual categories. Let us call the intentional correlates of such categories 'events'. Regardless of whether or not one assumes internal representations of the action sequence to operate on event concepts, linguistic representations of it do require segmentation into units that can be labeled as instances of *unwrapping a tea bag*, *pouring water into a cup*, and so on. A semantic typology of event segmentation is concerned with the conditions under which dynamic stimuli are broken down into instances of semantically distinct categories across languages. Events are not generally encoded by lexical items alone, but by syntactic constructions, such as verb phrases or clauses. As a result, the set of possible linguistic representations of a given event stimulus may not be enumerable, since its members may vary from one another along an indefinite number of choice points. The typology of event segmentation is therefore addressed here in terms of the *constraints* different

[1] Of course, not all events are perceivable. But perceivable events are a reasonable starting point. The scope of the problem is limited further below.

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



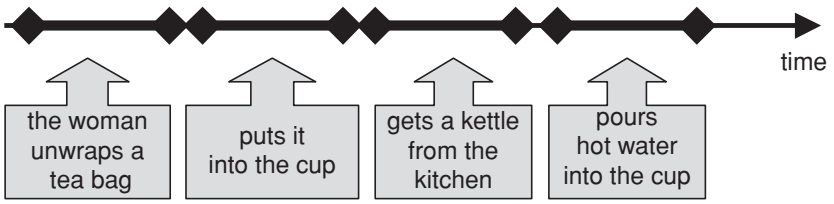


Figure 3.1 Event segmentation – an introductory example

languages impose on the segmentation of dynamic stimuli into semantic event categories. We argue that such constraints derive partly from lexicalization patterns (Talmy 1985) and partly from the availability of particular syntactic constructions.

Semantic typologies map the extensions of language-particular semantic categories on an 'etic grid', a possibility space created by a few independent notional dimensions in which every linguistic categorization in the domain under study can be located as a data point. These dimensions are the potential independent variables of the analysis. They are selected on the basis of evidence from prior research. The cells of the grid are then exhaustively encoded in sets of non-verbal stimuli and preferred descriptions and/or ranges of possible descriptions of these are collected in a typologically broadly varied sample of unrelated languages with multiple speakers per language, according to a standardized protocol. Etic grids are arguably a necessary prerequisite of cross-linguistic studies of linguistic categorization – at least as part of the implicit background assumptions of such studies. And in the interest of proper evaluation and critique of a study's protocol, laying out the grid explicitly is to be preferred. Etic grids bias the data collected on their basis, but they do not obscure this bias. For example, Levinson (2000) demonstrates that not all languages lexicalize color foci identifiable in terms of the two dimensions of the Munsell color chart – hue and brightness – employed as an etic grid in Berlin and Kay (1969). Yet, the demonstration is based primarily on data collected with the Munsell color chips.

The need for a manageable etic grid makes a typology of the segmentation of scenarios such as the action sequence depicted in fig. 3.1 rather ambitious. The research reported on in this chapter has focused on much simpler stimuli. The case study we present examines the encoding of causal chains across languages. Attention is specifically on stimuli in which a state change or location change is caused by one or more preceding events. Constraints on the linguistic representation of causal chains turn out to be sensitive to the number and types of 'sub-events' in the chain, as well as to the specific nature of the 'links', i.e. the causal relations among the sub-events. Event segmentation thus reflects the complexity of the semantic representations to be conveyed. At the elementary level at which the problem is analyzed here, complexity of the

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



internal structure of event representations – the number and types of sub-events to be encoded and the nature of the relations between them – emerges as the fundamental dimension on which constraints on event segmentation operate. Cut-off points along this dimension vary according to language-specific constraints. Language-specific patterns of event segmentation can be located on the complexity dimension in a way broadly similar to how language-particular color categories can be mapped on the hue–brightness matrix.

But first, a methodological issue requires some attention. If language-specific event categories are not simply delimited by the extensions of lexical event labels (again, chiefly verbs), then how are such categories to be identified? This problem is addressed in the following section.

## 2 The macro-event property

Pawley (1987) provides a study of event segmentation in Kalam, an East New Guinea Highlands language. Pawley compares Kalam and English in terms of how they segment event descriptions into 'conceptual events'. He defines 'conceptual event' as the meaning of a clause that contains a single 'event classifier', i.e. verb. The study finds striking differences between the two languages in the sets of possible conceptual events. In particular, Kalam lacks 'episodic' verbs, i.e., verbs that lexicalize script-level action sequences such as denoted by *make a cup of tea* as a summary description of the scenario in fig. 3.1 above. For instance, there is no simple verb that means 'hunt'. Instead, hunting activities are conventionally construed as sequences of four to six 'conceptual events', according to the schema in (1) (Pawley 1987: 344; the events in parentheses may or may not be mentioned):

| (1) | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| | (GO FORTH) | KILL GAME | BRING IT TO CAMP OR HOME | COOK IT | EAT IT | (RETURN HOME) |

(2) gives an example:

(2) *. . . mneb ak lgl mdek*
KAL land that having.come.about it:existed:DS
*kmn ak pak dad apl, ty ty*
game that kill carry having.come what what
*gl, adl ňbek . . .*
having.done having.cooked he:ate
'. . . when that land came into existence, people hunted game mammals [and cooked and ate them]' (Pawley 1987: 338)[2]

[2] Abbreviations in morpheme glosses include: 1 – 1st person; 3 – 3rd person; A – 'Set-A' cross-reference; ABL – ablative; ABS – absolutive; ACAUS – anticausative; ACC – accusative;

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



The difference in lexicalization between the two languages is obvious. The question is whether this amounts to a difference in what is semantically represented as an instance of an event category.[3] Events are encoded in language and cognition as having 'merological' (i.e., part-whole) structures. Parts and combinations of events are themselves conceptualized as instances of events (Casati and Varzi 1999; Krifka 1998; Zacks and Tversky 2001). So even if the hunting activity is broken down into a series of 'conceptual events' in (2), these still "add up" to a representation of hunting as an event. Moreover, different verbs and verb phrases are in different syntactic relations in (2). Some are more "tightly integrated" syntactically – i.e., more similar to simple sentences – than others. How do such syntactic differences affect the semantics and pragmatics of the event representation? Should one not assume, contrary to Pawley, that the relative syntactic complexity or simplicity of the expression has an impact on the complexity or simplicity of the semantic event representation it conveys? Givón (1991b) compares on-line and off-line descriptions of a video stimulus in four Papuan languages including Kalam, which heavily makes use of serial verb constructions and clause chaining, and in Tok Pisin (or Neo-Melanesian, the English-based creole used as a lingua franca in Papua New Guinea), which has few serial verb constructions and no chaining. He finds that pauses of a certain length are significantly less likely to occur inside serial verb constructions than elsewhere, regardless of language. From this he concludes (p. 120) that "serial verb constructions do not represent a different cognitive way of segmenting reality."

Pauses may not be a very reliable measure of event segmentation, either, since they are likely to reflect a host of factors in addition to semantics which may (although they need not) be independent of event segmentation (including phonology, syntax, and pragmatics; see Levelt 1989: 256–60, 385–7). But at any rate, Givón's study suggests that serial verb constructions in Kalam form tighter syntactic units than clause-chaining constructions and sequences of independent clauses. And it stands to reason that event segmentation is affected by such differences in syntactic packaging. Consider the examples in (3), which may be representations of the same stimulus event:

ALL – allative; APP – applied object; AUX – auxiliary; B – 'Set-B' cross-reference; CAL – calendrical; CAUS – causative; CL – classifier; CMP – completive; COM – comitative; CON – converb; D2 – indexical (distal/anaphoric); DEF – definite; DIM – diminutive; DIR – directional; DS – different subject; EXIST – existential/locative; EVID – evidential; F – feminine; FOC – focus; GEN – genitive; HESIT – hesitation; IMPF – imperfective; IN – inanimate; INC – incompletive; INST – instrumental; LOC – locative; M – masculine; NEG – negation; NOM – nominative; PAST – past; PERF – perfect; PL – plural; PRES – present; PRV – perfective; REL – relational; SPONT – spontaneous; SG – singular; TOP – topic.

[3] A complication is that (2) appears to have habitual or generic reference. So strictly speaking, it refers to an indefinite number of instances of an action sequence of the same kind.

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



(3) a. Floyd opened the door.
b. Floyd pushed the door open.
c. Floyd pushed the door and it opened.

The verb *push* in (3b) specifies how Floyd caused the door to open. In (3a), the causal sub-event leading to the door's opening is present in the semantic representation, but it is left unspecific – (3a) does not reveal what exactly Floyd did to cause the door to open. Example (3c), like (3b), explicitly refers to the pushing event and the opening event, and invites a strong implicature (a defeasible inference) to the effect that these are causally related. But in (3b), this causal relation is in fact entailed. And this does not appear to be the only semantic difference between (3b) and (3c). In (3b), the pushing event and the opening event are entailed to be in spatio-temporal contiguity. In (3c), these relations are again merely implicated.

The syntactic relation between event-encoding phrases reflects or encodes the semantic relation that is expressed to obtain between the events referred to. What is called for, then, is some measure that assesses how event segmentation is affected by the syntactic properties of the construction. Is there a way of telling whether (3b) or (3c) construe the pushing and the opening as parts of one event or as two separate events? No, not exactly. This is not possible, because, again, the sub-events of (3c) are at least *implicated* to form parts of a larger event, just as all the events referred to in a narrative of indefinite length may be construed to form part of a single event. The difference in event segmentation between (3b) and (3c) lies in the "tightness" of "packaging". And that is primarily a difference, not in *what* is expressed, but in *how* it is expressed – a difference in the *mapping* between form and meaning.[4] But there is way of assessing this mapping difference.

The defining ontological characteristic of events[5] is that they are individuated, not just in space (as "objects" or "things" are),[6] but also in time. Events "occupy" time intervals and have a beginning or an end in time – most typically, both – and duration. The *existence* or *history* of objects of course is time-bound as well; but whereas different "time slices" out of the course of an event individuate distinct parts of the event, it is not the case that parts of the history of an object define parts of the object. Therefore, it makes sense to assume that

[4] Again, there is a purely semantic difference between (3b) and (3c), but that difference does not concern event segmentation *per se*, but merely the distinction of what aspects of the event representation are entailed vs. implicated.

[5] As intentional correlates of event concepts and event expressions – no claims are made here concerning the existence of events in extralinguistic/extracognitive reality.

[6] There are, in fact, abstract objects and events whose spatial individuation is problematic – e.g. things like democracy, inflation, or poetry. But all events, even abstract ones, are individuated in time.

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



it is the temporal properties of a construction that provide the decisive clues to its event construal. Indeed, the three event descriptions in (3) differ crucially in this respect: (3c) admits distinct time-positional operators in the two verb phrases:

(3) c′. Floyd pushed the door and it opened *immediately* / *after a moment of breathless suspense*.

This is impossible in (3a–b). With these descriptions, any operator that defines a position in time or duration necessarily has scope over both sub-events:

(3) a′. Floyd opened the door *immediately* / *after a moment of breathless suspense*.
b′. Floyd pushed the door open *immediately* / *after a moment of breathless suspense*.

In both cases, the temporal operators express the temporal distance between the combination of the door's opening and Floyd's pushing (or the unspecified causal event in 3a′) and some reference point – not the distance between the pushing and the opening event, as in (3c′).

In precisely the sense that the pushing sub-event is not accessible to operators of temporal position or duration at the exclusion of the opening sub-event, and neither is the latter at the exclusion of the former, these sub-events are not semantically "individuated" in (3a–b), but are presented as parts of an event that in terms of the criteria of duration and location in time is unanalyzed. But this is quite clearly a *structural* property of (3a–b) – a mapping property of the basic clause structures of (3a–b) at the syntax–semantics interface. We may call this property the **macro-event property** (MEP), borrowing the term 'macro-event' from Talmy (2000a). For a more formal definition, see Bohnemeyer *et al*. (2007). For present purposes, the characterization in (4) should do:

(4) **Macro-event property (MEP)**: An event-denoting construction has the MEP iff it combines only with those time-positional or durational operators that have scope over all sub-events it entails.[7,8]

In the remainder, the MEP serves as a heuristic: the encoding of complex causal chains across languages is examined with respect to the question as to which constructions involved in it have the MEP and which do not. Put differently, it is explored to what extent there is uniformity or variation in what parts

[7] It should go without saying that (4) is restricted to semantically and syntactically well-formed combinations of event-encoding constructions and temporal operators.

[8] The observation that differences in syntactic packaging result, aside from differences in the division of labor between semantics and pragmatics, in the differences in form-to-meaning mapping properties captured by the MEP, goes back to the Generative Semantics debate; see Fillmore (1972), Fodor (1970), and Wierzbicka (1980: 162–3).

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



of the stimuli are segmented as 'macro-events', i.e., encoded by expressions that have the MEP. Thus the MEP plays a role in these studies comparable to the role of the 'conceptual event' unit in Pawley's comparison of Kalam and English. The advantage of employing the MEP as the primary criterion in a typology of event segmentation is its sensitivity to the syntactic "packaging" of event reference. It is demonstrated in Bohnemeyer *et al*. (2007) that multi-verb constructions may have the MEP and there are mismatches between clause-hood and the MEP. Moreover, as discussed in detail in Bohnemeyer (2003) and Bohnemeyer *et al*. (2007), there are specific constraints on form-to-meaning mapping that emerge as operating, not on a particular unit of phrase structure, such as the clause or verb phrase, but on whatever construction has the MEP. An example is the bi-uniqueness constraint on the encoding of thematic roles (Bresnan 1982; Chomsky 1981; Fillmore 1968, *inter alia*). This indicates that the MEP is not merely an otherwise arbitrary property that happens to be quite suitable for the purposes of a typology of event segmentation. Indeed, the MEP appears to play a substantive role in constraining form-to-meaning mapping at the syntax–semantics interface.

## 3 Design of the study

The present study grew out of a larger project, an examination of the semantic typology of event segmentation in the domains of motion, causality, and transfer (or change of possession) undertaken by the members of the Event Representation Project at the Max Planck Institute for Psycholinguistics between 1999 and 2004. The study was conducted with a two-pronged design, combining a questionnaire and a video stimulus. The questionnaire – called Event Integration Questionnaire – consisted of a structured list of complex event scenarios represented in a semantic metalanguage, to be used, not in direct elicitation, but as a checklist – the researchers were to collect renditions of the questionnaire scenarios in the target languages by whatever technique seemed applicable, including with the help of the video stimulus (see Bohnemeyer 1999 for further details). The video stimulus – the Event Complexity (ECOM) clips – comprised seventy-four short animated videos representing complex events that involved a number of simple geometrical objects (circles, rectangles, triangles; see Bohnemeyer and Caelen 1999).[9,10] Both the questionnaire and the ECOM

[9] The researchers negotiated culturally appropriate interpretations of the objects and their motions with the consultants; e.g., Mayan speakers interpreted a triangle as a pyramid.

[10] Several contributors to the study worked, instead of or in addition to ECOM, with the real-video stimulus Staged Events, developed by M. van Staden, G. Senft, N. Enfield, and J. Bohnemeyer specifically for issues of event encoding in multi-verb constructions. Staged Events includes renditions of the ECOM scenarios featuring location change sequences, realized with a remote-controlled toy car moving around in a model landscape. See van Staden *et al*. (2001).

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



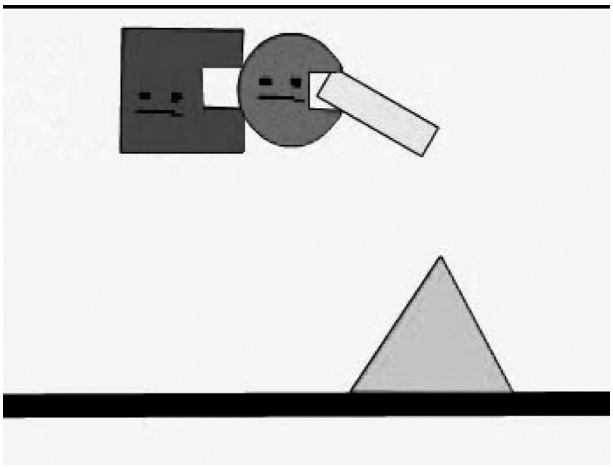

Figure 3.2 ECOM E7

clips were used to collect descriptions of complex stimulus events under two conditions: (a) the most natural descriptions of the various scenarios in the languages under investigation; and (b) the most "densely packaged" descriptions of the scenarios acceptable in the target languages, i.e., (roughly) those descriptions that made do with the smallest number of clauses while still entailing all relevant subevents (as prescribed in manuals accompanying the two tools). Results of this research in the motion domain, drawing on primary data from eighteen languages, are reported in Bohnemeyer (2003) and Bohnemeyer *et al*. (2007).

The data collected with the ECOM clips and the Event Integration Questionnaire turned out to be insufficient as the basis for a semantic typology of the segmentation of causal chains. This is because a surprising language-independent tendency manifested itself in the ECOM descriptions to leave causality largely to implicature. Consider the ECOM clip E7: a blue square bumps into a red circle, causing it to drop a yellow bar onto a green triangle, which breaks; see fig. 3.2.[11]

Regardless of language, descriptions of this clip are very similar, as far as the encoding of causal relations are concerned, to the English description in the previous sentence. Here are Dutch and Yukatek Mayan examples:

[11] Some of the characters of the ECOM clips were given facial expressions to motivate the idea of them controlling inanimate objects (instruments or themes in transfer scenarios).

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



(5) (. . .) komt een paarsig haakje, komt het beeldscherm binnen.
DUT comes a purple hook comes into the screen
Botst tegen een rood rondje aan met een geel staafje.
bumps against a red round thing with a yellow staff
Op het moment dat ie daar tegen aan botst
at the moment that it against it bumps
valt het gele staafjevan het rondje af
falls the yellow staff from the round thing off
en komt terecht op een groen driehoekje
and lands on a green triangle
dat in tweeën splitst . . .
which in two splits
'(. . .) a purple hook appears on the screen. Bumps into a red round thing with a yellow stick. The moment it bumps into it, the yellow stick falls off from the round thing and lands on the triangle, which splits in half (. . .)'

(6) (. . .) k-u=chíik-pah-al le=chan kwàadradro=o',
YUK IMPF-A.3=appear-SPONT-INC DEF=DIM square=D2
chich u=tàal=e', k-u=koh-ik
hard(B.3.SG) A.3=come(INC)=TOP IMPF-A.3=collide-INC(B.3.SG)
le=chan (. . .) sìirkulo=o', le=chan sìrkulo túun=o',
DEF=DIM circle=D2 DEF=DIM circle then=D2
k-u, óolbèey, estée, k-u=lúubul
IMPF-A.3 it.seems HESIT IMPF-A.3=fall-INC
hun-p'éel chan che'-il yàan (. . .) ti'=e',
one-CL.IN DIM wood-REL EXIST(B.3.SG) LOC(B.3.SG)=TOP
k-u=hats'-ik le=chan triàangulo=o',
IMPF-A.3=hit-INC(B.3.SG) DEF=DIM triangle=D2
k-u=káach-al.
IMPF-A.3=break/ACAUS-INC
'(. . .) the little square appears, it comes on hard, it bumps into the little (. . .) circle; the circle now, it, apparently, uhm, a little piece of wood that (. . .) [the circle] has falls, it hits the little triangle, [the triangle] breaks.'

Neither of the two descriptions contains a single causative light verb or a single caused-state-change verb, even though both languages have plenty of both. The Yukatek speaker actually employs an *anti*causative (or 'middle-voice') form of a caused-state-change verb (*kach* 'to break something') to refer to the final state change in the chain (the collapse of the triangle) – and this is quite typical. Clearly this phenomenon deserves further attention. For now, one reasonable interpretation of this phenomenon seems the following: Because the ECOM clips feature event sequences, the descriptions are in a narrative format (or 'genre'). Apparently, there is something of a conflict between narrating

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



events in the "main story line" and the encoding of causal links between these events. It seems that causality is either omitted or backgrounded in narratives (see, e.g., Lascarides 1992). And since in the case of the ECOM clips the causal information is already perfectly recoverable from the event information alone due to Gricean stereotypicality implicatures, speakers do not bother to background it, as that would feel redundant.

The lack of causative expressions in the ECOM descriptions forced us to develop a more targeted approach to the elicitation of causal chain descriptions. We assembled a new set of stimuli and designed an elicitation procedure that relies on two types of questions: first, questions as to why a certain event featured in a particular clip happened ('*Why*-questions'), and second, questions as to which participant caused the event ('*Who*-questions').[12] Researchers were instructed to ask these questions about as many of the events in each of the causal chains featured in the videos as seemed appropriate. In addition, they were asked to pay attention specifically to the first and last link in the causal chains, probing for an expression that would attribute the cause for the final state change (in the scenario in fig. 3.2, the breaking of the green triangle) to the event participant that sets off the whole chain (in fig. 3.2, the blue square). The researchers would do this by offering a range of possible constructions to the native speaker consultants and asking for the best choice. For example, in the case of the scenario in fig. 3.2, the range of possible expressions of the causal relation between the square's bumping into the circle and the triangle's breaking in English might look as follows:

(7) a. Did the square break the triangle?
b. Did the square make the triangle break?
c. Did the square cause the triangle to break?
d. Did the triangle break because of the square('s bumping into the circle)?

The same range of causative expressions was to be used in the *Who*-questions to the extent they were applicable:

(8) a. Who broke the triangle?
b. Who made the triangle break?
c. Who caused the triangle to break?

Where the researchers felt a need to avoid existential presupposition, they were to use the form in (9) first:

(9) a. Is there someone in this video who broke the triangle?
b. Is there someone in this video who made the triangle break?
c. Is there someone in this video who caused the triangle to break?

[12] The model for this approach is the '*where*-question' in the elicitation of locative expressions pioneered with the now famous BowPed stimulus (see Bowerman and Pederson ms.; Levinson, Meira, and the Language and Cognition Group 2003).

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



The report in the following section focuses specifically on the encoding of the causal relation between the "first" and "last" participants in the chain, elicited either with the *Who*-question or with the approach illustrated in (7) above.

We used a combination of twenty-one ECOM and eleven Staged Events clips.[13] The working title for this selection of stimulus items is **ECOM Causality Revisited (ECR)**.[14] The videos were selected so as to achieve a broad-based representation of various factors that might be conceived of as contributing to a given scenario's (language-specific) ranking on a scale of **directness** of causation (see Comrie 1981; Kemmer and Verhagen 1994; Shibatani 1976a; Talmy 1988; Verhagen and Kemmer 1997; *inter alia*). Directness of causation is the central dimension of the etic grid for this study. We assume that directness of causation breaks down into a number of independent factors. Specifically, directness varies in the ECR scenes along the following parameters:

**Mediation**: the number and type of 'links' in the causal chain. To simplify matters, the problem is reduced here to the number of event participants involved in the chain and the roles they play in it. Of course these roles are ultimately determined by the kind of event in which a participant is involved. Four role types are distinguished: 'causer' (CR) – the participant who sets off the causal chain; 'causee' (CE) – an animate "intermediate" participant who may or may not have some degree of control over the event (s)he is involved in; 'instrument' (IN) – an inanimate "intermediate" participant over which the CR/CE has complete control;[15] and 'affectee' (AF) – the participant undergoing the state change that marks the final link in the chain.[16,17] Four mediation types may be distinguished in the ECR clips along these lines: CR > AF (a causer directly effecting a change on an affectee without involvement of a causee or instrument); CR > IN > AF (a causer effecting a change on an affectee with

[13] Again, Staged Events is a real-video stimulus developed specifically for the study of event representation in multi-verb constructions. It includes real-video renditions of some ECOM scenarios, but also additional scenes not instantiated in ECOM. See van Staden *et al.* (2001).

[14] See Bohnemeyer and Majid (2002).

[15] It is implied here that causee and instrument are poles of a continuum. A type of role intermediate between the two that is of some relevance in the ECR scenarios is that of an inanimate object over which causer has no control, or insufficient control.

[16] In order to qualify as "intermediate", a participant has to (a) be acted upon (or causally affected) by the next "higher" participant in the chain and (b) act itself upon (or causally affect) the next "lower" participant.

[17] These labels designate roles as part of the etic grid (see section 2) of the study and thus should not be confused with semantic roles in language-specific causative constructions that encode any of the stimulus scenarios. For example, a particular linguistic representation may well choose to "emically" frame as a causer what is "etically" a causee. Note also that the terminology used in this study deviates from a convention often found in the literature (e.g., Kemmer and Verhagen 1994; Verhagen and Kemmer 1997) which refers to the final participant in the chain as 'causee' unless an intermediate causee is involved, and only in that case calls the final participant 'affectee'.

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



the help of some instrument); CR > CE > AF (a causer effecting a change on an affectee via a causee); and CR > CE > IN > AF (a causer effecting a change on an affectee via a causee and with the help of some instrument).[18]

**Contact**: this refers to spatio-temporal contiguity of the various events in the causal chain, or conversely, to the presence of spatial or temporal "gaps". For example, someone who hits a plate with a hammer affects it more directly than someone who merely hits the table on which the plate is placed. Similarly, the agent with the hammer may be viewed as less directly responsible for the plate's breaking if the breaking occurs not instantly, but only after some lapse of time.[19] Many of the ECR items featuring spatio-temporal gaps were modeled after the ballistic collision displays Albert Michotte used in his classic research on 'phenomenal causality' (see, e.g., Michotte and Thinès 1963). Since lack of spatio-temporal contiguity can affect any and every link in a causal chain, the set of combinatorial possibilities is large, and only a relatively small number of the possible combinations are actually instantiated in ECR.

**Force dynamics**: Talmy (1988) has argued that causation is conceptualized as a special type of interaction of (mechanical or metaphorical) forces. In the simplest case, a "stronger" 'agonist' "overpowers" a force-dynamically "weaker" 'antagonist', thus forcing a change of state against the antagonist's inherent tendency. Another pattern, which Talmy calls 'letting' or 'enabling', results when a stronger antagonist ends its impingement on a weaker agonist's inherent tendency towards change. Force-dynamic patterns, too, may influence the directness of causation. For instance, if someone drops a plate and it shatters upon hitting the floor, the person may be thought to cause the plate's breaking less directly than if (s)he had smashed the plate to pieces with a hammer. In the former case, gravity takes part of the blame, as it were. The number of force-dynamic patterns distinguished on Talmy's approach is large all by itself, and once again, there is no *a priori* reason why any of these patterns could not apply to any of the links of a causal chain. Only a small fraction of these possible combinations are realized in ECR. The stimuli include seven clips in

[18] There is in fact only one instance of CR > CE > IN > AF realized in ECR – namely in the ECOM E7 clip sketched in fig. 3.2 above. In a more complete stimulus set, one would obviously also want to include the options CR > IN > CE > AF and CR > IN > CE > IN > AF.

[19] Intuitively, the entire domain of *psych* causation – animate causees under the impact of external causers carrying out activities which are primarily internally caused (see Smith 1978) – is intimately tied to a lack of spatio-temporal contiguity. Not only do psych causation events not seem to involve contiguity, but many of the ECR stimuli that feature "gaps" are interpreted in language after language as involving psych causation. It should be stressed in this connection that none of the ECR stimuli were *designed* to show psych causation. The domain of ECR, from an etic perspective, is physical causation, albeit in some instances physical causation across a spatial gap or after a temporal delay.

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



which enabling-type dynamics obtain between causee and instrument, causer and affectee, or causee and affectee; all of these involve the pull of gravity. All other stimulus items feature exclusively causation-type dynamics.

Four languages were included in the study reported here: Ewe (studied by Essegbey), Japanese (Kita), Lao (Enfield), and Yukatek (Bohnemeyer). All languages were studied in the field, i.e., in Ghana, Japan, Laos, and Mexico. The number of speakers consulted ranged from three in the case of Lao to seven in the case of Yukatek, with four speakers for Japanese and six for Ewe.

## 4 The encoding of complex causal chains across languages

Examples of initial evidence for typological variation in the domain of causality can be found in Pawley's (1987) study of event segmentation in Kalam. Kalam has few verbs that lexicalize caused state change, which is commonly expressed by serial verb constructions and clause chains instead:

(10) kab añañ ap yap pkek, pagak ok
KAL stone glass come fall it:having:struck:DS it:broke that
'A stone fell and struck the glass and it broke' (Pawley 1987: 355)

Pawley notes that (10) is the closest Kalam equivalent of *The stone broke the glass*.

The four languages in our sample are typologically extremely diverse in terms of their resources for the encoding of causal chains. At the lexical level, Lao differs from the other three languages in that – not unlike Kalam – it has few simple verbs that lexicalize caused state change.[20] Not one of the thirty-two ECR scenes can be described in a single clause headed by (or projected from) a single transitive verb. The most simple construction used in reference to any of the ECR scenarios is a 'resultative' multi-verb construction (MVC) in which the first verb phrase may be transitive or intransitive and denotes the causing event, and the second verb phrase is intransitive and encodes the caused event:

(11) kuu$^{3}$ thup$^{1}$ còòk$^{5}$ tèèk$^{5}$
LAO 1.SG smash glass break
'I smashed the glass (and it) broke'

(12) *kuu$^{3}$ nùng$^{1}$ moong$^{2}$ thup$^{1}$ còòk$^{5}$ sòòng$^{3}$ moong$^{2}$ tèèk$^{5}$
LAO 1.SG one hour smash glass two hour break
'I smashed the glass at one (and it) broke at two'

As (12) illustrates, this construction has the macro-event property (MEP). Ewe has similar resultative MVCs, which likewise have the MEP. However, these

[20] There are transitive verbs that lexicalize destructive activities, such as *thup$^{1}$* 'crush'; however, these merely implicate, but do not entail, state change.

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



play a lesser role in the ECR data, since they compete with simple transitive verb clauses.[21]

(13) e-tutu-i do ɖa
EWE 3.SG-push-3.SG send away
'[The circle] pushed [the square] away'

The domain of such resultative MVCs is properly included in that of simple transitive verb clauses in Ewe. Resultative MVCs are by and large restricted to unmediated (CR > AF type) chains in which there is contact between causer and affectee at the time of change.

The core domain of simple transitive verb clauses in Ewe, Japanese, and Yukatek within the etic grid of this study is that of unmediated (CR > AF) spatio-temporally contiguous chains that involve only causation dynamics. The same may be said of the Lao 'resultative' multi-verb construction in (11). Mediation by an instrument has only a minor effect on the acceptability of the simplest construction; to be more precise, whether or not the presence of an instrument has an effect on the applicability of the simplest construction seems to depend mostly on whether it is indeed construed as an instrument. Mediation by a causee, however, renders the simplest construction squarely inapplicable in Japanese and Lao and disfavors it in Ewe and Yukatek.[22]

There is some variation in the extensional range of the simplest construction along the parameters of contact and force dynamics. Lack of contact strongly disfavors simple transitive verb clauses in Yukatek, but not in any of the other languages. Delays between impact and change – temporal discontinuities – disfavor simple transitive verb clauses in Japanese, while they have no effect by themselves in any of the other languages. Ewe simple transitive verb clauses become dispreferred only under the *combined* impact of lack of spatial and temporal contiguity. The Lao resultative MVC is strongly disfavored by enabling dynamics, which has no effect on the simplest construction in any of the other languages – as a matter of fact, enabling dynamics does not have a demonstrable effect on *any* construction in *any* language, aside from the Lao resultative MVC. Thus, it appears that simple transitive verb clauses in Yukatek are contact-sensitive, while their Japanese counterparts are timing-sensitive and their Ewe equivalents are contact-and-timing-sensitive (but in the sense that lapses of spatial and temporal contiguity only have an effect

[21] These Ewe constructions are known under a range of different labels in the literature; see, for instance, Ameka (2005a,b) for a recent summary.

[22] Ewe speakers readily describe chains mediated by causees with simple transitive verb clauses provided they interpret the causer as clearly intending the outcome. This would correspond to the condition under which Japanese speakers may apply morphological causatives to a chain mediated by a causee, as noted above; however, it seems that the threshold for attributing intention may be different in the two cultures.

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



where both types occur simultaneously). Lao resultative MVCs, in contrast, are force-dynamics-sensitive. Future research will need to clarify the source of these cross-linguistic differences. There are three possibilities here – none of which are mutually exclusive. It is possible that these differences are rooted in the semantic properties of the constructions in question. Or they may be pragmatic consequences of some formal (structural) properties. It is also very much conceivable that the extensional differences of the constructions reside in culture-specific nuances in the conceptualization of causality.

Japanese and Yukatek are grouped together against Ewe and Lao by the presence of causative morphology, i.e., affixal valence-changing or voice operations that license a causer argument semantically linked to the participant of a newly introduced causal subevent. In Yukatek, causative morphology is restricted to ('unaccusative') intransitive verb roots or stems that encode non-internally caused events (mostly state changes), such as *lúub* 'fall' in (14); see Bohnemeyer (2003) for details:

(14) t-u=lúub-s-ah
YUK PRV-A.3=fall-CAUS-CMP(B.3.SG)
úuch u=koh-ik le=x-ch'úup=o'
happen(B.3.SG) A.3=bump.into-INC(B.3.SG) DEF=F-female=D2
'[He] caused [the plate] to fall the way he bumped into the woman'

There is no such restriction in Japanese. Compare, for instance, (15), with the non-internally caused denominal verb *idoo-s* 'change location', and (16), with a transitive stem:

(15) akai maru-ga aoi shikaku-o oshi-te
JPN red circle-NOM blue square-ACC push-CON
idoo-s-ase-ta
change.of.location-do-CAUS-PAST
'The red circle made the square change its location by means of pushing it'

(16) onna-no hito-ga otoko-no hito-ni
JPN female-GEN person-NOM male-GEN person-LOC
osara-wo war-ase-ta
plate-ACC break-CAUS-PAST
'The woman made/let the man break the plate'

However, even though (16) encodes a causal chain involving three participants (a causer, causee, and affectee), no description of this type actually occurred in the ECR data. The reason for this conspicuous absence appears to be that the construction type (16) strongly implicates that the causer, the woman in (16), *intended* the caused event, in (16), the breaking of the plate. Intended outcome is not clearly applicable to any of the ECR clips that involve mediation by a

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



causee. Overall, morphological causatives play only a marginal role in the ECR data from either language, and in both data sets, the conditions under which morphological causatives are used are more or less identical to the conditions under which underived transitive verbs are used.[23]

Three of the four languages – all except for Japanese – have periphrastic constructions involving causative 'light' verbs. In all instances, these constructions have the macro-event property (MEP). In Ewe, *na* 'give', 'make' and *wɔ* 'do' are used as causative light verbs. The complement referring to the caused event may be intransitive or transitive:

(17) ŋutsu-a na (be) agba gbã
EWE man-DEF make that plate break
'The man made the plate break'

(18) nyɔnu-a ye na/wɔ-e (be) ŋutsu-a gbã agba
EWE woman-DEF FOC make/do-3.SG that man-DEF break plate
'It is the woman who made the man break the plate'

As (17)–(18) show, the complementizer *be* is optional in this construction. Example (19) illustrates the MEP:

(19) *ŋutsu-a na/wɔ-e etsɔ
EWE man-DEF make/do-3.SG yesterday
(be) ŋutsu-a gbã agba egbea
that man-DEF break plate today
'Yesterday the woman caused the man to break the plate today'

Periphrastic causative constructions in Lao employ *haj*$^{5}$ 'give', *hêt*$^{1}$ 'make', or a combination of both:

(20)[24] man$^{2}$ hêt$^{1}$ kèèw$^{4}$ tèek$^{5}$/sia$^{3}$
LAO 3 make glass break/be.lost
'He broke/lost the glass'

(21) man$^{2}$ hêt$^{1}$-haj$^{5}$ kuu$^{3}$ met$^{2}$ ngen$^{2}$ laaj$^{3}$
LAO 3 make-give 1 finish money much
'He caused me to lose a lot of money'

The semantic differences between these choices are not yet fully understood. There are no restrictions in terms of the transitivity of the complement. Thus, as in Ewe, periphrastic causative constructions can be used to encode chains that involve a causee, as in (21). Examples (22)–(23) show that these constructions have the MEP:

[23] Both Japanese and Yukatek also use compound verbs to encode causal chains. But again, these are used for the same types of scenarios – in terms of the distinction built into the etic grid of the study, as laid out above – as simple transitive verbs are.

[24] Superscript numbers indicate register tones in the Lao orthography used here.

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



(22) *man$^{2}$ nùng$^{1}$ moong$^{2}$ hêt$^{1}$ kèèw$^{4}$ sòòng$^{3}$ moong$^{2}$ tèèk$^{5}$/sia$^{3}$
LAO 3 one hour make glass two hour break/be.lost
'He at one broke/lost the glass at two'

(23) *man$^{2}$ nùng$^{1}$ moong$^{2}$ hêt$^{1}$-haj$^{5}$
LAO 3 one hour make-give
kuu$^{3}$ sòòng$^{3}$ moong$^{2}$ met$^{2}$ ngen$^{2}$ laaj$^{3}$
1.SG two hour finish money much
'He caused me at one to lose a lot of money at two'

In Yukatek, causative periphrases are formed with *mèet* 'do', 'make'. The complement may be intransitive (externally or internally caused) or transitive, and in the latter case, it may appear in the active, anticausative (or 'middle'), or passive voice.

(24) leti' le=chan tàabla=o'
YUK it DEF=DIM plank=D2
k-u=mèet-ik uy=op'-ik
IMPF-A.3=make-INC(B.3.SG) A.3=burst-INC(B.3.SG)
le=máak le=chan triàangulo y=éetel le=chan che'=o'
DEF=person DEF=DIM triangle A.3=COM DEF=DIM wood=D2
'That little plank [i.e. the blue square], it made the person [i.e. the red circle] burst the triangle with the stick'

The complement has to be transitive in order to permit the encoding of a causee. The causee is linked to the "subject"[25] of the embedded verb in the active voice or to an adjunct in the passive. Consider the contrast between (25), where the complement appears in the anticausative and the intermediate participant in the chain (the hammer) is construed as an instrument, marked by the comitative preposition *éetel* 'with', and (26), where the complement is in the passive and the hammer is construed as a causee, marked by the causal preposition *tuméen* 'by', 'because of':

(25) t-u=mèet-ah uy=óop'-ol
YUK PRV-A.3=make-CMP(B.3.SG) A.3=burst/ACAUS-INC
y=éetel le=máartiyo=o'
A.3=COM DEF=hammer=D2
'(S)he made it burst with the hammer'

[25] The 'A'-argument in the sense of Dixon (1994), or the 'actor'-argument in the parlance of Van Valin and LaPolla (1997). If there is a grammatical relation of subject in Yukatek – which is not obvious – then it is not consistently marked; see Bohnemeyer (2004) for details.

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



(26) t-u=mèet-ah uy=op'-a'l
YUK PRV-A.3=make-CMP(B.3.SG) A.3=burst-PASS.INC
tuméen le=máartiyo=o'.
CAUSE DEF=hammer=D2
'(S)he caused it to be burst by the hammer'

Example (27) illustrates the MEP:

(27) *Juanita=e' byèernes-ak=e' t-u=mèet-ah
YUK Juanita=TOP Friday-CAL=TOP PRV-A.3=make-CMP(B.3.SG)
u=mìis-t-ik u=nah-il Pedro sàabado
A.3=broom-APP-INC(B.3.SG) A.3=house-REL Pedro Saturday
'Juanita, last Friday, she made Pedro sweep her/his house on Saturday'

In all three languages that have causative periphrases, mediation by a causee strongly favors this construction type. Lack of spatio-temporal contiguity likewise makes periphrastic causatives the preferred choice over simpler constructions. In Lao – and only in Lao – enabling dynamics also favors periphrastic constructions.

Japanese lacks periphrastic causative constructions of the kind illustrated above, in first approximation, because it lacks a suitable causative light verb. Together with the restriction of synthetic (morphological) causatives to scenarios in which the caused event is clearly intended by the causer, the absence of causative light verb constructions imposes a set of constraints on the event segmentation of causal chains in Japanese that differs dramatically from that in the other three languages. The next least complex (or most densely packaged) alternative to simple transitive verb clauses in Japanese is a variety of constructions that employ 'converb' forms, i.e., subordinate verb forms morphologically marked for various semantic relations between the event or proposition expressed by the subordinate clause and the event or proposition referred to by the main clause (see Hasegawa 1996). Some of these constructions have the MEP, but most do not. Among the converb constructions featured in the ECR corpus, the only ones that have the MEP employ a *-te* converb:

(28) onna-no hito-ga osara-o teeburu-ni
JPN female-GEN person-NOM dish-ACC table-LOC
tataki+tsuke-te wat-ta
hit+attach-CON break-PAST
'The woman broke the dish by smashing it against the table'

(29) onna-no hito-ga hanmaa-o
JPN female-GEN person-NOM hammer-ACC
otoshi-te sara-o wat-ta
drop-CON dish-ACC break-PAST
'The woman broke the dish by dropping a hammer'

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



Here, the matrix clause encodes a causal chain involving causer and affectee, and the converb clause serves to further specify the causing event. The subjects of the two clauses must be coreferent. This construction does not permit the encoding of a causee. Its domain is largely coextensive with that of simple transitive verb clauses. It is slightly favored over plain transitive clauses by scenes that involve an instrument, especially under lack of contact between causer and affectee at the time of change. Examples (30)–(31) illustrate the MEP:

(30) *onna-no hito-ga osara-o teeburu-ni tataki+tsuke-te
JPN female-GEN person-NOM dish-ACC table-LOC hit+attach-CON
go-fun-go-ni wat-ta
five-minute-later-LOC break-PAST
'The woman broke the dish five minutes later [i.e., after smashing it] by smashing it against the table'

(31) *onna-no hito-ga hanmaa-o otoshi-te
JPN female-GEN person-NOM hammer-ACC drop-CON
go+fun+go-ni sara-o wat-ta.
five+minute+later-LOC dish-ACC break-PAST
'The woman broke the dish five minutes later [i.e., after dropping the hammer] by dropping a hammer'

This construction contrasts with the one illustrated in (32) which has the causal converb formative *-node*:

(32) te-de hageshiku teeburu-o tatai-ta-node
JPN hand-COM hard table-ACC hit-PAST-because
osara-ga ware-ta
plate-NOM break-PAST
'Because (someone) hit the table hard, the plate broke'

Here, the *-node* converb clause semanto-syntactically behaves very much like a causal adverbial clause in English – it has its own tense inflection and subject. (The two subjects are referentially disjoint – notice that *waru* `break' in (31) is transitive, but *wareru* 'break' in (32) is intransitive.) As (33) shows, this construction lacks the MEP:

(33) te-de hageshiku teeburu-o tatai-ta-node
JPN hand-COM hard table-ACC hit-PAST-because
go-fun-go-ni osara-ga ware-ta
five-minute-later-LOC plate-NOM break-PAST
'Because [someone] hit the table hard, the plate broke five minutes later'

A variety of other converb constructions occurred during ECR elicitation – these do not even entail a causal relation between causing and caused event, but merely implicate such a relation. The construction in (33) is semantically

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



Figure 3.3 Early and late frame of ECR 18

and in terms of its form-to-meaning mapping properties equivalent to causal connective constructions in the other three languages. Ewe, Lao, and Yukatek all possess such constructions. (34) is a Lao example:

(34) kuu$^{3}$ bòò$^{1}$-daj$^{\circ}$-paj$^{\circ}$-lin$^{5}$ tòòn$^{3}$ sòòng$^{3}$ moong$^{2}$
LAO 1.SG NEG-achieve-go-play period two hour
ñòòn$^{4}$ kuu$^{3}$ bò$^{\circ}$-mii$^{2}$ ngen$^{2}$ tòòn$^{3}$ nùng$^{1}$ moong$^{2}$
because 1.SG NEG-have money period one hour
'I didn't go out at two because I didn't have any money at one'

As the example shows, this construction does not have the MEP. However, no single ECR clip required a connective construction to express the causal relation between CR and affectee in Ewe, Lao, or Yukatek, or even elicited a connective construction as the preferred response. In contrast, exactly half of the ECR clips elicited a non-MEP converb construction as the preferred response type in Japanese, and in response to five of these sixteen scenes, a non-MEP converb construction was in fact the only option of encoding the causal relation between causer and affectee. Lack of contact and delays between cause and effect both favor non-MEP converb constructions in Japanese, and mediation by a causee leaves it as the only option of encoding a causal relation. This means that whereas all thirty-two ECR scenarios can be represented by single macro-event constructions in Ewe, Lao, and Yukatek, Japanese speakers prefer multiple macro-event constructions in half of the cases, and are left with multiple macro-event constructions as the only resource for representing causal chains mediated by a causee. Just as the study on motion event segmentation summarized in Bohnemeyer *et al*. (2007) unearthed profound and systematic cross-linguistic differences in the constraints imposed by language on the encoding of location change sequences, so our study reported on here has found dramatic systematic differences in the constraints different languages impose on the segmentation of causal chains.

The findings of the ECR study may be illustrated with a selection of scenes suitable to bring out the cross-linguistic variation that has been discovered. One stimulus item that is categorized as minimally complex across the four languages is ECR 18, as depicted in fig. 3.3. A red circle slides or rolls across the screen until it "hits" a stationary blue square. The two figures then travel on together in the same direction until they leave the screen. This is the type

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



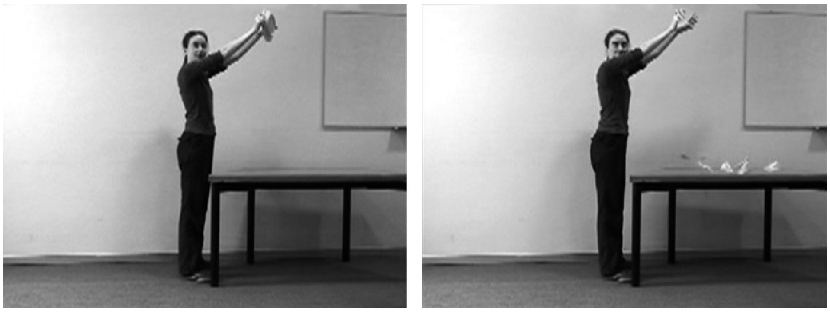

Figure 3.4 Early and late frame of ECR 5

of scene Michotte and Thinès (1963) dubbed 'entraining'. The causal relation between the motion of the circle and that of the square can be described by a simple transitive verb clause in Ewe, Japanese, and Yukatek. Speakers of all three languages also offered more complex descriptions, using a resultative multi-verb construction in the case of Ewe, a multi-macro-event converb construction in the case of Japanese, and a periphrastic causative construction in the case of Yukatek. Lao, however, requires minimally a resultative MVC to encode the causal relation in this scenario, as it lacks simple transitive verbs that could do the job. Now consider ECR 31. This is identical to ECR 18, except that the circle never actually "touches" the square. It stops at a short distance from the square, and after half a second or so both objects start to travel in the same direction in which the circle moved before, keeping the distance between them constant. This clip thus features a disruption of contiguity both in the spatial and the temporal domain. This still can be described by a simple transitive clause in Ewe and Japanese, although the preferred strategy in Japanese is now a non-MEP converb construction. Yukatek speakers, however, require a periphrastic causative construction to express the causal interaction between the circle and the square, and Lao speakers – although they *can* use a resultative MVC in reference to this clip – likewise prefer a causative periphrasis.

Next, consider ECR 5: a woman drops a plate onto a table, and the plate shatters; see fig. 3.4. In this scenario, the causer still inflicts a state change on the affectee relatively directly (with her bare hands, literally), except for the involvement of enabling dynamics – the role of gravity. For speakers of Ewe, Japanese, and Yukatek, it is perfectly acceptable in this case to say that the woman 'broke the plate', using a simple transitive verb clause (in Yukatek, a periphrastic description is considered equally appropriate, due presumably to the lack of contact between CR and affectee at the time of change). But in Lao, due to the special role force dynamics appears to play in determining the applicability of causative constructions in this language, not even a resultative

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



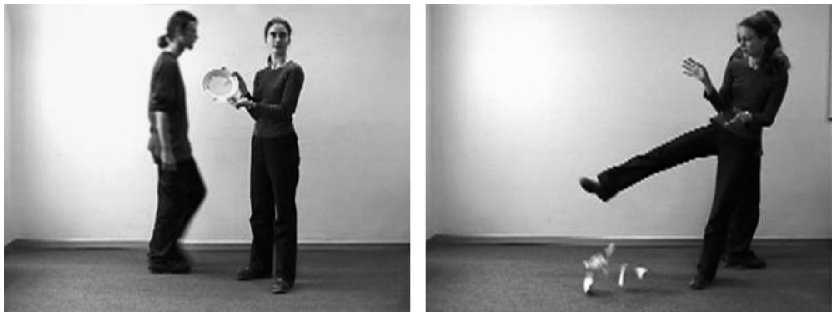

Figure 3.5 Early and late frame of ECR 23

MVC is acceptable in reference to this scene – one has to use a causative periphrasis.

ECR 5 contrasts minimally with ECR 32, in which the woman drops a hammer onto the plate, again with the effect of the plate shattering. The addition of an instrument to the causal chain has no effect on the Lao representation – again, periphrastic causative constructions are the only option. But in the other three languages, the level of formal complexity is revved up one notch to accommodate the increased conceptual complexity. In Ewe, a simple transitive verb clause is still the preferred response, but a resultative MVC emerges as an alternative. In Japanese, a converbial macro-event construction à la (29) above becomes the preferred response. And in Yukatek, a causative light verb construction now is preferred over a simple transitive clause.

Finally, consider ECR 23 – a man bumps into a woman, who is holding a plate. She drops the plate, and the plate hits the floor and shatters. To attribute the cause of the plate's breaking to the man, Ewe speakers may either use a simple transitive clause ('He (focus) broke the plate') or a periphrastic construction ('He (focus) made her break the plate') – both descriptions are considered equally acceptable. In Yukatek, a periphrastic representation is again preferred over a simple transitive verb clause in response to this scenario. In Lao, a periphrastic light verb construction is the only choice. And the Japanese consultants all opted for a non-MEP converb construction ('The plate broke because he tickled her') as the most densely packaged acceptable representation that encodes the causal relation between the man's action and the plate's breaking.

## 6 Summary and implications

This chapter has presented some building blocks of a semantic typology of event segmentation. The domain of event segmentation differs critically from

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



the domains of the classic studies of the cognitive anthropologists – terminologies for color categories, kinship relations, and ethnobiological classifications – in that it cannot be adequately captured in terms of lexicalization alone. Events are linguistically represented, not just by lexical items, but by morphosyntactic constructions and entire discourses. One could resort to comparing the semantic extension of particular types of event-denoting constructions – verb phrases, clauses, etc. – across languages. But this is likewise unsatisfactory, for a number of reasons. There *is* no single construction type that is uniquely dedicated to the encoding of events – so how would one generalize across construction types? And without such generalizations, how is one to capture the impact that differences in the availability of certain constructions have on event segmentation in particular languages?

The proposal that has been advanced here is to abstract away from individual construction types to a *property* of construction types that describes their behavior at the syntax–semantics interface, in such a fashion as to directly determine the semantic event representations a construction type is compatible with. This property is the *macro-event property* (MEP). Since events are critically individuated by their temporal properties – their boundaries, duration, and "location" (order and distance) relative to other events or times – the MEP registers the compatibility of event-denoting constructions with operators that modify or specify these temporal properties. For a construction to have the MEP means for it to "package" an event representation so "tightly" as to render its proper sub-events inaccessible to those temporal operators that might individuate them. By analyzing event descriptions into the constructions they consist of and evaluating these in terms of the MEP, the MEP provides a criterion of event segmentation that is sensitive to both lexicalization and morphosyntactic packaging, and that is readily applicable cross-linguistically irrespectively of the particular type of event-denoting construction.

The application of the MEP to the semantic typology of event segmentation in causal chains in the study presented here has indeed confirmed that event segmentation, in terms of the constraints imposed by individual languages on the information about an event that can be packaged into certain constructions, as assessed via the MEP, is a function of the interaction of lexicalization patterns and the availability of morphosyntactic constructions. Among the four languages considered, a split emerges between Japanese and the other three languages in the encoding of chains that involve causees (animate intermediate participants with some amount of control over the sub-event they are immediately involved in): Japanese requires the use of multiple macro-event expressions to encode the causal relation between the initial causer and the final state change (at least unless the outcome is construed as clearly intended by the causer), whereas the other three languages permit the encoding of these types of scenarios in single macro-event expressions. The reason is a combination of

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



lexical and syntactic factors: Japanese lacks causative light verb constructions, and in fact lacks the requisite causative light verbs; and causative morphology is restricted to intended outcomes. (There are other interesting lexical and morphosyntactic differences in this domain between the four languages, such as the lack of caused state change verbs in Lao; but these happen not to affect the segmentation of chains among macro-event expressions.)

The case study reported on here has confirmed Givón's (1991b) claim that lexicalization alone is not an adequate measure of event segmentation. Yet, at the same time it has also fully confirmed the extent of the cross-linguistic variation in event segmentation argued for on the basis of the contrast in lexicalization between English and Kalam in Pawley (1987). This variation has been found to occur not just in lexicalization, but to "project upwards" into constraints on syntactic packaging, given particular kinds of interaction between lexical and morphosyntactic factors.

The ultimate question raised by the cross-linguistic variation in event segmentation we have found is, naturally, that of its implications for the language–cognition interface. A classic relativist view would be that internal cognitive event representations vary with linguistic constraints. A cognitive universalist, in contrast, would argue that linguistic event representations are "supported" by the same internal cognitive representations, regardless of how many macro-event expressions they may have to be segmented into depending on the constraints imposed by particular languages. We do not at this point have any evidence that bears on this debate. We would, however, like to point out that the macro-event property (MEP) is neither a purely syntactic nor a purely semantic property. Consider (35)–(36):

(35) a. Sally went from Nijmegen to Arnhem.
b. Sally left Nijmegen and then went to Arnhem.

(36) a. Floyd pushed the door shut.
b. Floyd pushed the door and it shut.

Examples (35a) and (36a) have the MEP, and (35b) and (36b) do not. But (35a) and (35b) convey the same information, and so do (36a) and (36b). The difference between (35a) and (35b) is mostly in the division of labor between semantics and pragmatics, and the same holds for (36a) and (36b).[26] In particular, the causal relation between Floyd's pushing and the door's closing is entailed in (36a), but merely implicated in (36b). But in order to implicate the same scenario as in (35a), a speaker uttering (35b) must have the same scenario "in mind" in some sense. The same holds again with respect to (36a–b). Barring

[26] In addition, there are differences in lexical aspect: (35a) and (36a) are accomplishments, (35b) is a sequence of two achievements, and (36b) is a sequence of an achievement, or activity and an achievement.

P1: SJT Trim: 228mm × 152mm Top: 12.548mm Gutter: 18.98mm
CUUK1097-03 CUUK1097/Bohnemeyer ISBN: 978 0 521 89834 8 July 6, 2010 10:18



further evidence, we tentatively conclude that the difference between (35a) and (35b) lies not so much in the underlying mental representations, but in the mapping between these conceptual representations and syntax. The same holds for (36a) and (36b).

The MEP is a form-to-meaning mapping property of event-denoting constructions at the syntax–semantics interface. The evidence from interface constraints such as the bi-unique mapping of thematic relations being sensitive to the MEP (as discussed in Bohnemeyer *et al*. 2007) suggests that the MEP is built into the design of human language itself. But we see no reason to assume that for the MEP to operate in language, it needs to be supported by an ontological category of "macro-events" in internal cognition.