Bengt Sigurd & Jordan Zlatev

Motion Event Typology Meets Computational Modelling

Abstract

Speaking involves a process of selecting words and constructions which correspond to the thoughts that one wishes to express. Different languages place different constraints on this selection process, corresponding to their typological characteristics. In this paper we show how Swedish, French and Thai preferentially express motion events differently and suggest a theoretical framework with grounds in cognitive science to account for these differences. We proceed to describe an experimental computer program, *Verbalizer*, which implements aspects of our framework, showing how a single conceptual representation of a motion event can be realized in multiple ways, both within and between the languages. Our research indicates how typological and computational linguistics can benefit from one another.

1. Introduction

Motion events involving change of location like rocks falling, horses running and children going to school are universal phenomena, but their encoding in language differs—both within and between languages. Extensive work in spatial semantics over the past two decades (for summary, see e.g. Zlatev in press), implies that there are (at least) three major types of languages, depending on the preferred way of expressing motion events. These three types can be illustrated with examples (1–3) from Swedish, French and Thai in which the speaker recalls the event of swimming across a river:

(1)	*Jag*	*simma-de*	*över*	(*flod-en*)	
	1SG	swim-PAST	across	river-DEF	

(2)	*J'*	*ai*	*traversé*	*le*	*fleuve*	(*à la nage*)
	1SG	AUX	cross/PAST	DEF	river	swimming

A Man of Measure
Festschrift in Honour of Fred Karlsson, pp. 460–471

(3)	*chán*	*wâinásam*	*khâam*	*mææenáam*
	1SG	swim	cross	river

In Swedish, Manner of motion is usually expressed by the verb, while Path is expressed by a particle/preposition and the Landmark expression is optional. In French the main verb typically encodes Path, while Manner is expressed by an optional adverbial. Finally, Thai typically codes both Manner and Path with two different verbs in a serial verb construction. In each one of the three languages (and their corresponding types), it is also possible to use a number of other utterances permitted by the grammar and lexicon expressing more or less the same event, but the *preferred* patterns differ.

Computational linguistics has not been much concerned with matters of linguistic typology, but it is clear that defining not only the possible but also the preferred pattern for expressing a proposition should be of high relevance, in particular for machine translation. Conversely, typology could benefit from a more explicit specification of the dimensions along which languages differ, teasing apart general cognitive, semantic and syntactic factors, and as is well-known, computational linguistics can provide efficient tools for formalization and experimentation. In this article, we intend to show how cognitive typology and computational linguistics could mutually benefit from a closer relationship.

2. Theoretical framework: Vision, space and language

A good deal of work within cognitive science and related fields has been dedicated to explaining the relationship between language and (visual) perception, e.g. Miller & Johnson-Laird (1976), Harnad (1990), Regier (1996), Gärdenfors (2000). If there is any consensus among these very different approaches, it is that language does not "map" onto perceptual experience directly, but rather onto intermediary concepts.

In his analysis of human vision for the purpose of developing robotic systems, Kopp (2003) shows that in order to approach the complexity of visual cognition (at least) three different levels of visual processing are required: (a) early vision involving simple *features* and templates, (b) intermediary vision involving "*visual agents*", as in the Pandemonium Model of Selfridge (1973) eventually giving rise to 3-dimensional perception and (c) high-level vision, involving *routines* from different but interrelated systems: *What, Where, How* and *When*. The relation between the different levels is not a matter of “bottom-up” processing, but a highly interactive one. Furthermore, Kopp points out the need of a still-higher level of abstraction, which he calls *meta-schematic,* “which integrates information from view independent representations, and thus is not concerned with basic perceptual processes” (Kopp 2003: 148). The meta-schematic level is therefore not really perceptual but conceptual. The distinction is similar to the one between perceptual and conceptual structures emphasized from a developmental perspective by Mandler (2004), who calls the child’s first conceptual structures “image schemas”, inspired from work in Cognitive Linguistics. The latter term is, however, highly ambiguous (see Hampe 2005), so we will refer to basic conceptual structures underlying spatial and temporal cognition as *meta-schemas*. They can be seen to be in part based on the routines of high-level vision, but are at the same time shaped by language itself (Bowerman 1996). Table 1 shows the basic set of meta-schemas involved in spatial semantics, some of their possible *values,* i.e. the kind of phenomena that can be categorized under them, and the perceptual systems that motivate them.

Conceptual representations (CRs) of motion events can thus be seen as consisting of configurations of meta-schemas and their values. Thus the CR that corresponds to the sentences (1)–(3) can be specified as shown in (4), which can also be expressed as “I crossed the river by swimming over it”, even though this is not a very idiomatic rendition in English.

Meta-schema	**Values**	**Perceptual system**
Trajector (TR)	Open-ended	*What*
Landmark(s) (LM)	Open-ended	*What*
Region (of Landmark)	IN, ON, UNDER, ABOVE…	*Where*
Frame of Reference (FoR)	OBJECT-CENTERED VIEWPOINT-CENTERED GEOCENTRIC	*Where*
Motion	YES/NO	*When*
Path	BEGINNING, MIDDLE, END	*Where*, *When*
Manner • Gait • Speed • Body part	 RUN, WALK, FLY… SLOW, MID, FAST… FEET, WINGS, KNEES …	*How*

Table 1. Meta schemas, values and perceptual systems

(4) Motion: YES, Manner: SWIM, Path: MIDDLE
TR: EGO, LM: RIVER, Region: ABOVE, FoR: OBJECT-CENTERED

Given the assumptions of our framework, we can predict that languages may differ on the basis of the following factors:

- **Granularity**: which values are relevant, e.g. Korean makes the contrast TIGHT-FIT/LOOSE-FIT (Bowerman 1996) in the meta-schema Region, while English does not. With respect to the schema Path, Archi (Dagestanian) makes the difference between END-TOWARDS (allative case) and END (terminative case).

- **Conflation patterns**: more than one meta-schema expressed by a single form-class, e.g. *enter* → Motion:YES + Path:END + Region:IN

- **Distribution patterns**: conversely, it is possible for a single meta-schema to be expressed in several form-classes, often redundantly, as in Swedish, where the combination of verb particle and preposition *in* + *i* ← Region:IN

- **Constraints**: Since languages are highly conventional symbolic systems, it is customary for them to include various semantic and grammatical constraints on "well-formedness" which can not be derived from general cognitive or processing factors. One such constraint is the so-called *Boundary Crossing* constraint, specifying that Manner verbs cannot be used when the Trajector crosses a boundary (Slobin 1996). The Romance languages in general obey this constraint, but the languages of the other types do not.

3. Applying the framework to motion event typology

One way to display the differences between the three ways of expressing motion events, as exemplified by Swedish, French and Thai, is to show their predominant patterns of mapping meta-schemas and expressions, as shown in Table 2.

Mapping pattern	*Type 1 (Swedish)*	*Type 2 (French)*	*Type 3 (Thai)*
Conflation	V→Motion+Manner Part→Path+Region Prep→ Path+Region	V→Motion+Path+ Region	V→Motion+ Manner V→Motion+Path+ Region V→Motion+Deixis
Distribution	Part+Prep← Region	V+Prep←Region	V+Prep←Region V+V(+V)←Motion

Table 2. Conflation and distribution patterns within the mapping of meta-schemas and form-classes in the three language types (exemplified by Swedish, French and Thai).

In terms of conflation patterns, French and Swedish differ in terms of whether Path and Region are conflated with Motion in the verb root (e.g. *traverser*), or in a "satellite" (e.g. *över*)—as in the original proposal of Talmy (1985). Thai displays a third type in which a motion verb can either conflate Manner, Path+Region or Deixis, and it is possible to combine all three (in this order) in a serial verb construction. The distribution patterns in the three languages differ as well: Region is coded in the particle and preposition in Swedish (when the two are distinct, e.g. *in i*), but in the verb and preposition in French and Thai. Due to its serial verb nature, Thai also has the possibility of distributing the expression of Motion over the different verbs.

Furthermore, there are constructional differences, having to do with the preferred grammatical patterns in the three languages. Focusing on *intransitive translocation*, i.e. motion where the Trajector itself is the grammatical subject, and the motion event involves change of location, we can specify the major construction types for the three languages as shown in (5), (6) and (7). The notation specifies the linear order of form-classes—in this case in a completely "flat" structure, but if necessary, hierarchical structure can be easily added—and the meta-schemas expressed by them. The French pattern in (6) contains not one, but two constructions, depending on whether there is boundary crossing or not.

(5) *Type1:* NP:TR V:Motion+Manner Part:Region Prep:Region NP:LM

(6) *Type2:* NP:TR V:Motion+Manner Prep:Region NP:LM
(if no Boundary crossing)
NP:TR V:Motion+Path+Region Prep:Region NP:L Adv:Manner

(7) *Type3:* NP:TR V:Motion+Manner V:Motion+Path+Region V:Motion+Deixis Prep:Region NP:LM

While we are aware that this level of abstraction glosses over many details and is ultimately insufficient, one advantage of expressing the typological characteristics so formally is that it allows a fairly straightforward computational implementation, as described in the next section.

4. Computational implementation: *Verbalizer*

The computer program *Verbalizer,* written in Prolog (LPAprolog), takes as input a conceptual representation (CR), implemented as an unordered list of *features*, i.e. schema-value pairs, and delivers a phonological representation (PR), along with a functional representation (FR) of the sentence. On a general level, Verbalizer can be said to simulate the process of human sentence production, with the CR corresponding to the speaker's *informative intention*, the PR to the produced utterance, and FR to a mediating link. The CR corresponding to the sentences (1-3) and given in (4) is implemented as in (8), (adding the meta-schema Time, but omitting Frame of Reference).

(8)
```
[[tr, ego], motion, [manner, swim], [region, above],
 [path, middle], [lm, [def, river]], [time, past]]
```

The PR is constructed successively by processing the items in the original CR, and in the process building incrementally a FR of the sentence consisting of functional roles such as subj(S), predicate(P), obj(O), adv(A), etc. which is necessary for controlling word order and agreement. The processing of the features of the CR and the building of the FR and PR is handled by a set of predicates called *combos* of the general form `combo(A,B)`. Each combo predicate implements a particular conflation or distribution pattern, by taking a list of features (A) as input, and looking for particular features to map onto particular words-forms. The remainder of the input CR is given back as B. For example, the combo predicate generating the English word-form `entered` would look for the features: `motion`, `[region,in]`, `[path,end]` and `[time,past]` in the CR, and if these are found they would be subtracted from the transitional CR, while `entered` is added to the PR, and `pred(entered)` to the FR. The combo predicates are used recursively until all features of the CR are processed.

The following example shows how the CR in (8) in processed in 4 steps, building English FPs and PRs in the process. Since the CR is empty after Step 4, the output PR is generated as an English verbalization of the CR.

```
Step 1 CR: [motion, [manner, swim], [region, above],
           [path, middle],[lm,[def, river], [time, past]]
       FR: [subj (I)]
       PR: [I]

Step 2 CR: [[region, above], [path, middle], [lm, [def,
           river]]]
       FR: [subj (I), pred (swam)]
       PR: [I, swam]

Step 3 CR: [[lm, [def, river]]
       FR: [subj (I), pred(swam), prep(across)]
       PR: [I, swam, across]

Step 4 CR: [ ]
       FR: [subj (I), pred (swam), prep (across), obj ([the,
           river])]
       PR: [I, swam, across, the, river]
```

However, at the same time as the features of the CR are being "subtracted" during generation, the original CR is maintained in a special format to be referred to when necessary, e.g. when verb agreement requires information of the grammatical features of the subject, or when a distribution rule (such as Part + Prep ← Region) requires access to a feature that has already been removed from the transitional CRs. Since the combo predicates involve (unidirectional) mappings between CR, FR and PR, operating on the level of grammatical constructions, our system may be considered as an implementation of a version of Construction Grammar (Goldberg 1995).

We have written combo predicates to handle intransitive motion event expressions in English and Swedish (belonging to Type 1), French (Type 2) and Thai (Type 3), and in the remainder of this article will illustrate how the program can give rise to type-specific verbalizations, and at the same time capture certain generalizations. The English sentence in (9) is not the only English verbalization produced for the CR in (8). Rather, using the inbuilt backtracking mechanism of Prolog, Verbalizer finds alternative renditions, such as those in (10) and (11).

(9) *I swam across the river.*

(10) *I crossed the river by swimming.*

(11) *I moved across the river by swimming.*

The reason that (9) is generated prior to (10) and (11) is technically due to the *ordering* of the combo predicates: in particular, those conflating the meta-schema Manner have been ordered before those conflating the meta-schema Path. At the same time, (9) is a more idiomatic rendition of the motion event than any of the alternatives, at least in English. We thus venture the following generalization, predicting the preferred verbalization of a motion event.

> ***Generalization for Manner verbalization:*** *If a CR contains the meta-schema Manner, conflate this in the first Motion verb, unless prohibited.*

As can be seen in the process of generating the preferred Thai rendition of (8), this generalization holds perfectly for Thai, where if there is a Manner verb, it always *heads* the verb series (Zlatev 2003). Step 2 for Thai shown below involves the conflation of the schemas Motion and Manner in the verb *wâináam* ('swim'). Step 3 verbalizes Path (along with Motion, which is absent in the transitional CR, and therefore derived from the original CR). Finally, Step 4 produces a complete PR, even though the feature (schema) Time still remains in the CR, since as many East-Asian languages, Thai lacks tense.

Step 1
```
CR: [motion, [manner, swim], [region, above],
     [path, middle], [lm, [def, river]], [time,
     past]]
FR: [subj (chan3)]
PR: [chan3]
```

Step 2
```
CR: [[region, above], [path, middle], [lm, [def,
     river]], [time, past]]
FR: [subj (chan3), pred1 (wai2naam3)]
PR: [chan3, wai2naam3]
```

```
Step 3 CR: [[lm, [def, river]], [time, past]]
       FR: [subj (chan3), pred1 (wai2naam3), pred2
       (khaam4)]
       PR: [chan3, wai2naam3, khaam4]
```

```
Step 4 CR: [[time, past]]
       FR: [subj (chan3), pred1 (wai2naam3), pred2
       (khaam4) obj (mae4naam3)]
       PR: [chan3, wai2naam3, khaam4, mae4naam3]
```

Even for French, the combo predicates verbalizing Manner are ordered before those for Path, and if there is no crossing of boundaries, a Manner verb will be produced first, following the generalization, as in (12).

(12)	*La*	*boule*	*a*	*roulé*	*vers*	*l'*	*arbre*
	DEF	ball	AUX	roll/PAST	toward	DEF	tree

'The ball rolled toward the tree'

However, this possibility is blocked if there is boundary crossing, implemented in Verbalizer by a constraint specifying that:

> IF the CR includes the feature Path:MIDDLE,
> OR
> the feature Path:BEGINNING or Path:END AND Region:IN or Region:OUT,
> THEN do not conflate Motion and Manner in V.

The CR in (8) contains `[path, middle]`, hence the program cannot proceed by adding the verb *nager* ('swim') to the PR, but instead uses the verb *traverser* ('cross') conflating the features `motion`, `[region, above]`, `[path, middle]` and `[time, past]` in the expression *ai traversé* in Step 2. The feature `[manner, swim]` needs to be verbalized by an adverbial in Step 3.

```
Step 1 CR: [motion, [manner, swim], [region, above],
           [path, middle], [lm, [def, river], [time, past]]
       FR: [subj (je)]
       PR: [je]
```

```
Step 2 CR: [[manner, swim], [lm, [def, river]]]
       FR: [subj (je), pred ([ai, traversé])]
       PR: [je, ai, traversé]
```

Step 3
```
CR: [[lm, [def, river]]]
FR: [subj (je), pred ([ai, traversé]), adv ([à,
    la, nage])]
PR: [je, ai, traversé, [à, la, nage]]
```

Step 4
```
CR: [ ]
FR: [subj (je), pred ([ai, traversé]), adv ([à, la,
    nage]), obj ([le, fleuve])]
PR: [je, ai, traversé, [à, la, nage], le, fleuve]
```

Additionally, late rules which we will not go into involving the mapping FR-PR, change the word order to place the object *fleuve* before the adverb and an elision rule contracts the words *je* and *ai* into *j'ai,* producing the preferred rendition, given above in (2).

To summarize, the program *Verbalizer* offers a simple model which captures explicitly some (possible) universals and dimensions of variance within and between languages. The universals include the meta-schemas, the types of mapping patterns, and the general preference for encoding Manner in the main verb. Variations involve the specific values of the meta-schemas (above all for Region), the specific verbalization patterns, general constraints such *Boundary Crossing*, as well as detailed rules involving e.g. word order. To capture both cognitive motivations and language-specific conventions, (at least) three different representational levels were required: Conceptual, Functional and Phonological.

5. Summary and conclusions

In this article we have briefly outlined a theory of spatial semantics attempting to account for both universal patterns and cross-linguistic variation in the verbalization of motion events. The theory is based on work in cognitive science and follows previous research by the second author (Zlatev 2003). Then we proceeded to show how parts of this theory can be successfully implemented in a Prolog program, *Verbalizer*, using the experience and even some of the programs developed by the first author (Sigurd 1982). While our work is still highly exploratory and tentative, we hope to have shown how cognitive-typological and computational linguistics can benefit from at least some cross-fertilization. The need for explicit formalization of concepts, constraints and processes provoked by the computational approach made clear the following three points of possible theoretical significance:

- In mapping between a conceptual and a phonological representation, an additional grammatical ("functional") representation is required.

- It is possible to account for motion event typology in terms of differences in conflation and distribution patterns, operating on a single conceptual representation (contra strong versions of the Sapir-Whorf hypothesis, though language-specific patterns may still have cognitive effects).

- There appears to be a general cognitive motivation, possibly related to saliency, for encoding Manner in the clausal head (the main, or first verb), unless this is overridden by conventionalized language-specific constraints.

On the other hand, by exploring the ways of encoding typological patterns and constraints, the experimental program *Verbalizer* may have practical applications for natural language processing (NLP) systems seeking more general solutions to defining the "natural" modes of expression in different languages, and specifically for practical systems commenting on scenes where objects move in different environments.

References

Bowerman, Melissa (1996) Learning how to structure space for language—A cross-linguistic perspective. In Paul Bloom, Mary A. Peterson, Lynn Nadel & Merril F. Garret (eds.) *Language and Space*, pp. 385–436. Cambridge, MA: The MIT Press.

Goldberg, Adele E. (1995) *Constructions: A Construction Grammar approach to argument structure*. Chicago, IL: University of Chicago Press.

Gärdenfors, Peter (2000) *Conceptual Spaces: The geometry of thought.* Cambidge, MA: The MIT Press.

Hampe, Beate (ed.) (2005) *From Perception to Meaning: Image schemas in Cognitive Linguistics*. Cognitive Linguistics Research 29. Berlin & New York, NY: Mouton de Gruyter.

Harnad, Stevan (1990) The symbol grounding problem. *Physica D* 42: 335–346.

Kopp, Lars (2003) *Natural Vision for Artificial Systems: Active vision and thought*. PhD Dissertation, Lund University. Lund University Cognitive Studies 100. Lund: Lund University Cognitive Studies.

Mandler, Jean (2004) *The Foundations of Mind: Origins of conceptual thought*. Oxford Series in Cognitive Development. Oxford & New York, NY: Oxford University Press.

Miller, George A. & Johnson-Laird, Philip N. (1976) *Language and Perception*. Cambridge, MA: Belknap Press of Harvard University Press.

Regier, Terry (1996) *The Human Semantic Potential: Spatial language and constrained connectionism*. Neural Network Modeling and Connectionism Series. Cambridge, MA: The MIT Press.

Selfridge, Oliver G. 1973. Pandemonium: A paradigm for learning. In Leonard Uhr (ed.) *Pattern Recognition, Learning, and Thought: Computer-programmed models of higher mental processes*. New York, NY: Prentice-Hall.

Sigurd, Bengt (1982) Commentator: a computer model of verbal production. *Linguistics* 20: 611–632.

Slobin, Dan (1996) Two ways to travel. Verbs of motion in English and Spanish. In Masayoshi Shibatani & Sandra Thomson (eds.) *Grammatical Constructions: Their form and meaning*, pp. 195–220. Oxford & New York, NY: Oxford University Press

Talmy, Leonard (1985) Lexicalization patterns: semantic structure in lexical forms. In Timothy Shopen (ed.) *Language Typology and Syntactic Description*, Volume III: *Grammatical categories and the lexicon*, pp. 57–149. Cambridge: Cambridge University Press.

Zlatev, Jordan (2003) Holistic spatial semantics of Thai. In Eugene H. Casad & Gary B. Palmer (eds.) *Cognitive Linguistics and Non-Indo-European Languages*, pp. 305–336. Cognitive Linguistics Research 18. Berlin & New York, NY: Mouton de Gruyter.

—— (in press.) Semantics of spatial expressions. In Keith Brown (ed.) *Encyclopedia of Language and Linguistics*. 2nd edition. Oxford: Elsevier.

Contact information:

Bengt Sigurd
Department of Linguistics
Center for Languages
and Literature,
Lund University
Box 201
SE-221 00 Lund
Bengt(dot)Sigurd(at)ling(dot)lu(dot)se

Jordan Zlatev
Department of Linguistics
Center for Languages
and Literature,
Lund University
Box 201
SE-221 00 Lund
Jordan(dot)Zlatev(at)ling(dot)lu(dot)se