Relating ERP-Effects to Theories of Belief Update and Combining Systems

Ralf Naumann

Institut für Sprache und Information
Universität Düsseldorf
Germany

Abstract. A major challenge for any attempt at combining linguistics with empirical neurophysiological data from brain research, like ERPs for example, is to give an answer to the question of how the gap between psycholinguistic and formal models of specific aspects of language on the one hand and the neural architecture underlying neurophysiological measures on the other can be bridged (Baggio et al. 2010). An interesting research program based on the extension of the event calculus in Van Lambalgen & Hamm 2005 was launched in Baggio et al (2007, 2008, 2010) where it was shown how a correlation between two known ERP-effects, the N400 and the LAN, and semantic phenomena like the progressive can be established. In this paper we will present an alternative formal theory which is based on the technique of combining systems and in which the dynamics of information change is separated from the more static aspects of knowledge representation. Using this multi-layered architecture, we hypothesize that the LAN is related to the process of updating a discourse model in the light of new information about changes in the world whereas the N400 concerns the more static aspects of situation models, e.g. the relation between events and persistent objects.[1]

Keywords: ERP, situation models, combining systems, aspect, belief revision

1 Introduction

A major obstacle for combining neuroscience and linguistics are the diverging ways in which the meaning and understanding of linguistic expressions (words) are defined. In theories of grounded cognition the meaning of a word is linked to (or even defined in terms of) past experiences and encounters with objects of the given kind. For example, the meaning of verbs like *kick* or *lick*, which refer to actions involving leg or mouth movements, is linked or identified with somatotopic activation patterns in corresponding motor cortices. Defining meaning in this way has the advantage of directly grounding the understanding and comprehension of language in the empirical evidence available to speakers.

---

[1] The research was supported by the German Science Foundation (DFG) funding the Collaborative Research Center 991 at the University of Düsseldorf.

In addition to past experiences with objects of the given kind a second source of learning the meaning are distributional properties abstracted from the linguistic contexts in which the word occurs. This way of defining the meaning of words fits well into the conception of viewing the primary function of language as a set of processing instructions on how to arrive at a situation model for the state of affairs referred to by a sentence or a discourse (Zwaan & Radvansky 1998).

This way of analyzing the meaning of verbs is almost orthogonal to the way they are analyzed in formal semantic theories. There verbs are classified e.g. as activities (*run* or *lick*), accomplishments (*write*) or semelfactives (*kick*). These classifications are based on properties of thematic relations like Theme and/or the notion of a nucleus structure. E.g. distinguishing *incremental* from *constant* Themes makes it possible to define accomplishment (*write*) and activity (*push*) verbs. The notion of a nucleus structure was introduced in Moens & Steedman (1988). These structures are built in terms of three different parts: *preparatory process*, *culmination* and *consequent state*. Verbs differ with respect to the type of nucleus structure assigned to them. For example, accomplishments verbs like *build* have nuclei structures consisting of all three possible parts, whereas the nucleus structure of an activity verb like *run* consists only of a preparatory process.

One way of combining the two paradigms is the following. As emphasized in Zwaan & Radvansky (1998:162), a situation model can be seen as a *multidimensional* system consisting of various dimensions like space, time, causation and objects. The various dimensions are linked to each other by particular relations guiding the flow of information between those dimensions. The different formal semantic theories can be taken as modeling a particular dimension of such a multidimensional system. For example, a nucleus structure defines the dynamic meaning component of a verb, i.e. the way an event of a given type (say a writing) evolves (or occurs) in space and time, and therefore the dimensions space and time in a situation model. The object dimension and its relation to the dynamic dimension is defined in terms of thematic relations like Actor or Theme. However, there are at least two missing links. First, the aspect of combining the different dimensions with each other (the flow of information) is missing in those semantic theories. Here I am going to propose using the framework of combining systems to fill this gap. Second, semantic theories must be related to neurophysiological and behavioural data from brain and cognitive science. In particular, the decompositions used in those theories (properties of thematic relations, the various constituents of a nucleus structure) don't seem to be directly related to the empirical evidence used by speakers to grasp the meaning of lexical items. A first step to bridge this gap has been undertaken in Van Lambalgen & Hamm (2005) where the meaning of a sentence is seen as an algorithm (or a procedure) to construct a scenario (or, formally, a logic program) from the sentence. However, even in this approach it remains unclear how the proposed decomposition clauses are related to the empirical evidence about meanings available to comprehenders. The strategy proposed here to overcome this problem is based on the following empirical considerations. Events occur in time. This occurrence can be subdivided into different stages as described by a nucleus structure. Each stage of the development is characterized by (i) a subset of the objects involved in the event and (ii) particular activities and/or properties of those objects that are undertaken or affected during that stage. Given information about the type of event (verb), the stage the event is in (aspectual and temporal markers) and the role (thematic relation) of the object in the event, it is possible to infer (based on empirical knowledge of events of that type) the corresponding activities and/or properties and to built up the temporal dimension of the situation model in terms of them.

Two important questions that arise are: (i) How are situation models computed in the brain? and (ii) What is the relation between the various ERP components and the various dimensions of a situation model? Given that a situation model is multidimensional, certain effects may be related to particular dimensions and not to others. If this is true, Hagoort's (2003) thesis that those components are *domain-specific* must be extended to also include some form of *dimension-specificity*.

In the following section the experimental results of Baggio et al. (2008) are presented. In the third section an alternative framework is developed and compared to that of Baggio et al.

2 The Progressive and left anterior negativity

Using the formal semantic theory developed in Van Lambalgen & Hamm (2005) and Hamm & Van Lambalgen (2003) and based on ERP experiments, Baggio et al. (2008, 2010) present a unified account of the N400 and the LAN (left anterior negativity) by relating it to different semantic phenomena like the progressive and coercion. Similarly to other approaches, Baggio et al. link the N400 to the processing of semantic phenomena at the level of semantic (word) composition like priming (semantic relatedness) and the unification of variables. According to the analysis of the progressive in Hamm & Van Lambalgen (2003), the meaning of a telic VP like *write a letter* is an algorithm (called a 'scenario'), implemented as a particular type of logic program from which the expression's reference, a minimal model of the scenario, can be derived via resolution (unification). For telic VPs such a scenario can be seen as a (simplified) plan to reach a particular goal. In the case of writing a letter this plan basically consists of the goal (the finished letter) and an activity to attain this goal, here writing (as opposed to say dictating). Since such a minimal model does not contain any information about obstacles (possibly) terminating the writing, it is possible to derive the obtaining of this goal at some later time using additional general constraints like the law of inertia. This conclusion concerning the existence of a goal state is defeasible. Now consider the sentences in (1).

(1) a. The girl was writing a letter when her friend spilled coffee on the paper.
    b. The girl was writing a letter when her friend spilled coffee on the tablecloth.

First, it was correctly predicted that a higher N400 effect occurred for *tablecloth* compared to *paper*. Baggio et al. explain this by the fact that in the context of a writing event (described in the main clause) the noun *paper* is semantically more expected as referring to the location where the letter comes into existence than the noun *tablecloth*. As a result, semantic unification is easier (proceeds more smoothly) for *paper* compared to *tablecloth*. Second, they (correctly) expected a higher LAN for the VP *spill coffee on the paper* compared to the VP with 'paper' replaced by *tablecloth*. This is explained by the fact that in the former case the minimal model derived for the main clause (MC) has to be recomputed due to the fact that the goal state (the completed letter) must be suppressed since the subordinate clause (SC) contains information about a terminating event. When this part of the sentence is processed by a speaker there is therefore a higher processing load which is neurophysiologically reflected in the higher sustained anterior negativity. By contrast, for *tablecloth* the minimal model is simply extended in a monotonic way. Finally, they correctly expected a correlation between higher LAN-amplitudes and speakers asserting that the letter was not finished (probe sentence for 1a).

3 An alternative approach

Following Zwaan & Radvansky (1998), we assume that the successful understanding of a sentence consists in the ability to construct a situation model for the situation (or state of affairs) described by it. A situation model is a multidimensional system in which various dimensions are combined with each other. This multidimensional character will be modeled by using the technique of combining systems (Blackburn & de Rijke 1997, Finger & Gabbay 1992). A dimension is modeled as an ontology together with a family of n-ary relations. In the present context three such ontologies are distinguished: time points, eventualities and persistent objects, which are defined as follows: (i) Object structure $\underline{\mathbf{O}} =< O, \{A_a\}_{a \in Attr} >$ with O a (non-empty) set of persistent objects (like tables and chairs) and each A_a is an *n-ary* relation on O; (ii) Eventuality structure $\underline{\mathbf{E}} =< E, \{P_v\}_{v \in Verb} >$ where E is a (non-empty) set of event tokens (or event occurrences) and each P_v is a unary relation on E, which is an event type and its elements are event occurrences or event tokens like runnings ($v = run$) or writings ($v = write$) and (iii) $\underline{\mathbf{T}} =< T, <_T>$ is a flow of time with T a (non-empty) set of time points and $<_T$ a linear and atomic ordering.

The definition of the flow of information (i.e. the connections) between the various dimensions is based on the following considerations. Each event token e occurs on a particular interval and is related to persistent objects from $\underline{O}$ that bear a particular relation to it like Actor or Theme. The first relation between $\underline{\mathbf{T}}$ and $\underline{\mathbf{E}}$ is defined by the function τ from E to $\wp(T)$, which assigns to each $e \in E$ its *run-time* and which is required to be a convex set. In terms of τ the functions PP, Cul and CS are defined, mapping an event e to the interval corresponding to the PP, Cul (a singleton) and CS, respectively. This connection is called the

temporal-causal dimension of a situation model. The relation between $\underline{\mathbf{E}}$ and $\underline{\mathbf{O}}$ is modeled by a set $\{R_{tr}\}_{tr \in TR}$ of (functional) thematic relations and is called the *static* dimension of a situation model. Events and actions bring about changes with respect to time-dependent properties of those persistent objects participating in them (i.e. which bear one of the thematic relations defined for the event/action). By way of illustration consider an event of a girl writing a letter. Such an event involves a goal (the completed letter) and an activity (writing) to bring about this goal. Since the writing involves an actor, it will be assumed that there is a state the writer is in when doing the writing. This state will be encoded by $state_{writing}(x)$, and be called a *state-fluent.* Formulas like *completed(y)*, encoding the goal, are called *goal formulas.* During the occurrence of the event of writing the formulas are evaluated in a characteristic way. Whereas $state_{writing}(girl)$ is true during the whole event (except, possibly, at the beginning and the end point), *completed(letter)* is only true at the end point of the event and, in addition, after the event occurred. This relation is captured as follows. Let $Attr_v$ be the restriction of *Attr* to those state-fluents and goal formulas related to event tokens from P_v. Then Z is a function that assigns to each event e and time point $t \in \tau(e)$ an object structure $\underline{O'} =< O', \{A'_{a'_e}\}_{a' \in Attr_v} >$ where O' is a subset of the objects involved in e and $\{A'_{a'_e}\}_{a' \in Attr_v}$ is the restriction of $\{A'_{a'}\}_{a' \in Attr_v}$ to O'. This structure is called the *dynamic submodel of e at t.* In terms of Z a function Z^* is defined that assigns to an event e a sequence of pairs (t, M_t) that describes the evolution of the event e.
How is the function Z computed in the brain? Events evolve (develop) in time. This evolution can be subdivided into different stages as described, e.g. by a nucleus structure. Each stage of the development is characterized by (a) a subset of the objects involved in the event and (b) particular activities and/or properties of those objects that are undertaken or affected during that stage. So given information about (i) the type of the event (supplied by the verb), (ii) the stage the event is in (supplied by aspectual and temporal markers like progressive) and (iii) the role of the object (extracted from e.g. morphological information and/or word order), it is possible to retrieve, using empirically grounded knowledge about events and objects of those kinds in that particular role, appropriate activities and/or properties of the corresponding object for that stage or those stages that have already occurred up to the stage of the event described by the sentence. In our framework the above relationship is captured by axioms like those in (2) for *write* (see the end of the paper for reasons why the consequents are given in a very schematic way).

(2) a. $Actor(e) = x \wedge PP(e) = i \wedge P_{write}(e) \rightarrow state_{writing}(x)(i)$
b. $Theme(e) = y \wedge PP(e) = i \wedge P_{write}(e) \rightarrow state_{of\_creation}(y)(i)$

c. $Theme(e) = y \wedge Cul(e) = \{t\} \wedge P_{write}(e) \rightarrow completed(y)(\{t\})$

d. $Theme(e) = y \wedge CS(e) = i' \wedge P_{write}(e) \rightarrow completed(y)(i')$

In these axioms information about the static dimension (thematic relation) and the temporal-causal dimension (second conjunct in the antecedent) together with type information about the event is used to infer information about the dynamic dimension. For a progressive sentence, only the first two axioms (2a) and (2b) are used, since the event is described as ongoing. By contrast, for a sentence in the simple past, the third axiom is used too and for a sentence in the (present) perfect, the antecedents of all four axioms are satisfied.

Empirical evidence for axioms of this kind comes from experiments that show that aspect and other grammatical devices influence not only the way an object is simulated but also which objects are focused (Madden & Therriault 2008, Ferretti et al. 2009, Bergen & Wheeler 2010). For example, in *John is closing the drawer* (ongoing event) the mental simulation focuses on the act of pushing the drawer undertaken by the Actor John whereas in *John has closed the drawer* the simulation is focused on the goal state and therefore on the Theme in that state, i.e. the *closed* drawer. The way an object is simulated is therefore closely related to its role in the event and to its properties affected by the event. On this view, decompositional predicates like those used in the consequents of the axioms in (2) are directly related to the way objects are simulated in the brain while computing a situation model whereas in the antecedents only information is used that can be directly extracted from the surface structure of the sentence.

An important aspect that is missing so far is non-determinism. If an event token occurs, it is in general not uniquely determined how the occurrence continues. For example, if a girl is writing a letter, she may finish it or not. Formally this non-determinism is modeled in the following way. Each pair (t, M_t) is related to a set of (linearly) ordered sequences of pairs $(t', M_{t'})$ such that (i) each sequence is a possible way of how the event occurring at t evolves after t and (ii) this set is ordered by a relation $\leq_t$, which is required to be at least reflexive and transitive. When taking both conditions together, one gets, for a given t, a computation tree (with root t) where each outgoing arc is assigned a weight, giving the degree of expectancy that this transition will occur (though these weights are *not* represented in the framework sketched above). This ordering reflects the expectancy or the priming a comprehender uses when processing a sentences with that verb. One basis of such an ordering is world knowledge about events of the given type, say writings. Another source is linguistic knowledge, in particular knowledge about the contexts in which the verb preferably occurs, e.g. whether the verb can be modified by temporal adverbials like *in x time* or *for x time* ('>' means that this context is more primed).

(3) a. John wrote the letter in an hour > John wrote the letter for an hour.
    b. John ran for ten minutes > John ran in ten minutes.

Verbs which prime *in*-adverbials compared to *for*-adverbials are normally related to goals and/or results. E.g. writing usually involves a particular object that is intended to be brought into existence, like a letter or an article for instance. By contrast, verbs priming *for*-adverbials are usually used to describe events that have no intended goal like running for example. From this it does not follow that *write* cannot be used without a goal, as examples like *John wrote* show, and *run* can be used as involving a goal: *Bill ran a mile/to the station.* However, these contexts are less primed. Posible orderings (primings) for *write* and *run* (or *push*) are: (i) *write*: PP Cul CS $>$ PP and (ii) *run*: PP $>$ PP Cul (CS).

How is a situation model computed in the framework sketched above? It is assumed that this computation is based on the following two principles: processing (i) proceeds on a word-by-word basis and (ii) is done in accordance with Hagoort's Immediacy Hypothesis (IA): all sorts of information available to the comprehender is immediately used in parallel in order to arrive at a meaningful interpretation.

When a word is processed, information is added at different dimensions. This information will in general be underspecified. In particular, this will be the case if information is added to a dimension that is the combination of basic dimensions like the static or the dynamic dimension because then information from *all* dimensions involved is necessary to arrive at fully specified information. For instance, when processing *wrote* in *The girl wrote a letter*, one gets the following updates: (i) type information: the event is a writing (basic dimension $\underline{E}$); (ii) unification $Actor(e) = girl$ (static dimension) since the subject has already been processed. By contrast, information about the Theme is still underspecified: $Theme(e) = y$; (iii) 'triggering' the corresponding axioms for the nucleus structure (dynamic dimension). In the last section it will be shown how the data from Baggio et al. can be explained in our framework.

3.1 The girl and the coffee: an alternative interpretation

According to the Baggio et al. interpretation, the higher LAN is explained as follows. A formula that is true in the minimal model after processing the MC has to be retracted when this model is updated with the information contained in the (disabling) SC (revision of the model, non-monotonic update). This is formulated in the following hypothesis.

> *The Non-Monotonicity Hypothesis (NMH)*
> A higher LAN is expected when the current model has to be recomputed (revised) if new information is added that is incompatible with information in the current model. The amplitude of the LAN depends on the extent to which the model has to be recomputed.

One very simple method of calculating the measure of change to a model (or database) consists in counting the number of formulas that have to be revised (see Katsuno & Mendelzon 1992). Consider again the disabling condition (1a). The spilling event causally interferes with the ongoing event of writing referred to in the MC, bringing it (at least temporarily) to an end. Let this information be encoded by the formula $clip(e_{spill}, e_{write})$, which is added to the temporal-causal dimension of the current model. Integrating this information in the model does not effect the valuation of a formula in the dynamic submodel at t_{spill}. However, the girl is no longer simulated as writing a letter after t_{spill} so that the formula $state_{writing}(girl)$ will *not* be an element of the dynamic submodel at times $t' > t_{spill}$. Similarly, the writing event is no longer described as occurring and the letter is no longer taken as coming into existence, yielding a total of three revisions. Of course, more sophisticated methods may be necessary.

The Non-Monotonicity Hypothesis not only applies to cases involving a goal state but to any non-monotonic change affecting the current model, in particular scenarios like those in (4) where a goal state is not part of the most expected nucleus structure, but the SC contains a disabling condition.

(4) The girl was pushing a cart towards the barn when she was hit by a car.

In (4) no goal state is involved because *push a cart* is atelic and the PP does not introduce a culmination. Therefore, no higher LAN is expected to occur according to the Baggio et al. analysis. By contrast, applying the NMH, the hitting terminates the pushing, so that the girl is no longer simulated as being in a pushing state. The difference between (1a) and (4) consists in the fact that for the latter the number of revision is expected to be lower because no object is coming into existence so that there are only two revisions. Thus, the amplitude of the LAN for (1a) is expected to be higher compared to that for (4) using the simple counting criterion from above.

In our approach, the following, alternative explanation for the examples in (1) and (4) is possible. After processing the MC, the event is described as ongoing. For an accomplishment verb like *write*, this means that different ways of how the event continues are still open: the writing (PP) is continued until the culmination is reached (most primed); (ii) the PP is continued but the culmination is not reached due to an interruption, leaving the eventual attainment of the goal open and (iii) the writing is definitely terminated. Thus, no *unique* nucleus structure has been determined up to that point, in particular it is *not* determined whether the culmination or goal state (e.g. a completed letter) is eventually attained or not. When a disabling condition is encountered, the ordering on the possible continuations is changed. Several ways of how a new ordering is constructed are possible. At one extreme, the current most expected continuation is simply discarded. This will likely be the case when the MC is continued by *when she got a heart attack* or *when she died* (case (iii) above). An alternative reordering simply switches the current ordering between the different continuations so that the currently most expected ordering is still available but only demoted. This is likely to be the case for the example (1a) and case (ii) above because it is still possible that the girl (re)continues writing the letter on another sheet of paper. Changing the ordering of the possible continuations has the effect of changing the most expected scenario. This is formulated in the hypothesis below.

> *Reordering of Nuclei Structures Hypothesis (RH)*
> A higher LAN is expected whenever the ordering on the set of continuations at a point t of the temporal-causal dimension is changed due to a disabling condition.

On the RH, a reordering is also expected for (4). Recall that a verb like *push* is associated with a nucleus structure of type PP Cul CS, though this structure is less primed. Encountering the disabling condition in the SC of (4), this continuation has to be discarded. How is the cost of reordering the set of nuclei structures calculated? A very simple way of calculating the cost of a reordering operation consists in counting the number of reorderings. For example, completely discarding the most expected nucleus structure should be more costly than discarding a less primed nucleus structure. An answer to this question depends both on the exact way of how the ordering is defined and, of course, on experimental findings related to processing sentences in the brain. According to the RH, the higher LAN in the Baggio et al. data is not due to a change in the valuation of (atomic) formulas. In the disabling condition the goal state is never part of the model because it is never asserted to have been attained. This state is, however, part of the most expected nucleus structure (scenario) and therefore expresses the generalization (expectation) that e.g. writing a letter normally leads to a completed letter. The higher LAN is rather an effect of a change wrt. the expectation of how the situation will dynamically evolve. A higher LAN can therefore be taken as the dynamic counterpart of the N400. Whereas the N400 is related to expectancy at the *static* dimension of a situation model, the LAN is related to its *dynamic* dimension.

What empirical evidence can be used to decide between the Non-Monotonicity Hypothesis and the Reordering Hypothesis (of course, they need not exclude each other)? In order to give a tentative answer to this question, we first have to look at the relation between formal models of the event calculus and neural networks in the framework used by Baggio et al. In the variant of the event calculus developed in Van Lambalgen & Hamm (2005) the scenario associated with a clause (or sentence) has a minimal model. Minimal models can be regarded as the stable states of associated neural networks. Computing the minimal model for a clause, say the MC in the examples in (1), corresponds to the neural network settling into a stable state. Processing the next clause, say the SC in (1), the network settles into another stable state. The difference between (1a) and (1b), then, is that the transition from the stable state the network is in after processing the MC to the stable state after processing the SC is more costly in the disabling situation (1a) compared to the neutral situation (1b) because in the former but not in the latter case activation patterns across some units related to the goal state that were previously active have to be readjusted in the sense that the weights (or connection strengths) between these units are changed. This explanation of the relation between formal model and neural networks is based on two assumptions: (i) there is one global neural network that goes from one stable state to another during processing the different clauses of a sentence or discourse and (ii) the amplitude of the LAN is dependent on the extent to which connection strengths have to be readjusted.

In our multidimensional framework each nucleus structure associated with a verb has a minimal model for the dynamic dimension of the situation model. If the ordering on the nuclei structures is changed due to additional information (a different scenario becomes most expected), the connection strengths in the corresponding neural network have to be readjusted. It is this readjustment that triggers the higher LAN. On this view there is both a minimal model for the dynamic dimension of the SC and the MC (i.e. there are two separate minimal models, one for each dynamic dimension). At this point we apply a distinction introduced in Zwaan & Radvansky (1998). They distinguish the *integrated* and the *current* model. Whereas the integrated model is the (global) model built up by the previous discourse, the current model is the model computed for the sentence that is being processed. After the current model has been computed, the integrated model is updated with this current model, resulting in a new integrated model. The important point is that if the minimal model is redefined, this has no effect on the valuation of formulas at the dynamic dimension in the current model because the valuation only concerns state-fluents and goal formulas. For example, if the MC describes an event as ongoing, only axioms like those in (2a) and (2b) are triggered. Their consequents remain true, independently of any disabling condition which is added while processing the SC. For the consequents of those axioms which are related to the Cul and the CS, no decision is possible because no information about the corresponding state is available. As a consequence, neither the consequent nor its negation are part of the current model.

On this view the deactivation of units simulating the girl as being in a writing state is not the cause of a higher LAN. First, no formulas have to be retracted since the girl is *not* simulated as *not* being writing. More important, however, is the following argument. In a recent ERP study Willems et al. (2009) found a neural dissociation between action word understanding and mental imagery. Participants were presented with verbs related to hand actions (e.g. *throw*) and nonmanual actions (e.g. *kneel*). They either read these verbs (lexical decision task LD) or actively imagined performing the actions referred to by these verbs (imagery task IM). During the LD task effector-specific activation in premotor but not in primary cortex was elicited whereas in the IM task effector-specific activation in both types of cortices was found. The authors explain these findings by making a distinction between *implicit mental simulation*, which is related to prediction, and *explicit mental imagery*, which is reflective involving a covert enactment of the corresponding action.

Relating this distinction to our framework, one gets that computing (building up) the dynamic dimension of a siuation model corresponds to action verb understanding and therefore to implicit motor simulation in premotor (as opposed to primary motor) cortex. At this level (action) verbs are classified according to their associated ordered sets of nuclei structures. Such a classification is quite different from a classification in terms of explicit mental imagery. For example although *write a letter* and *grasp a needle* have identical (or at least very similar) nuclei structures, their corresponding mental imagery widely differ. Thus, although *write a letter* and *grasp a needle* differ with respect to their goal formulas (say, *completed(letter)* and *be_grasped(needle)*), these differences only show up when it comes to explicit mental imagery. And even for a verb like *grasp*, the mental imagery depends on the kind of object to be grasped: *grasp a needle* vs. *grasp a barbell.* Conversely, although *stroke* and *write* both (normally) involve hand movement, they are related to different ordered sets of nuclei structures. What is important for action verb understanding, therefore, is the set of ordered nuclei structures, i.e., the different ways an event can dynamically develop. On this view the consequents in dynamic axioms like those in (2) can be seen as amodal and highly abstract, symbolic descriptions of the concrete modal motor imagery that get replaced when implicit simulation ends and explicit imagery begins (see Willems et al. 2009:2397). In our framework those abstract descriptions are the consequents in dynamic axioms, which correspond to transitions in dynamic submodels. When simulation ends and imagery begins, an operation like transition refinement (see e.g. Blackburn & de Rijke 1997) takes place which replaces these transitions with (or cues them to) a modal motor imagery in primary motor cortex.

References

1. Baggio, G. et al. (2008). Computing and recomputing discourse models, an ERP study, Journal of Memory and Language 59, 36-53.
2. Baggio, G. et al. (2010). Coercion and compositionality, Journal of Cognitive Neuroscience 22, 2131-2140.
3. Baggio, G. and M. van Lambalgen (2007). The processing consequences of the imperfective paradox, Journal of Semantics 24, 307-330.
4. Bergen, B. & K. Wheeler (2010). Grammatical aspect and mental simulation. Brain & Language 112, 150–158.
5. Blackburn, P. and M. de Rijke (1997). Zooming in, zooming out, Journal of Logic, Language, and Information, 6, 5–31.
6. Ferretti, T. R., H. Rohde, A. Kehler and M. Crutchley (2009). Verb aspect, event structure, and coreferential processing. Journal of Memory and Language, 61, 191-205.
7. Finger, M. and D. Gabbay (1992). Adding a temporal dimension to a logic system. J. of Logic Language and Information, 1, 203-233.
8. Hamm, F. and M. Van Lambalgen (2003). Event calculus, nominalisation, and the progressive, Linguistics and Philosophy 26, 381–458.
9. Katsuno, H. and A. Mendelzon (1991). On the difference between updating a knowledge database and revising it. KR-91, 387–394, Cambridge.
10. Madden, C. and D. Therriault (2009). Verb aspect and perceptual stimulations, The Quarterly Journal of Experimental Psychology, 62:7, 1294-1303.
11. Moens, M. and M. Steedman (1988). Temporal ontology and temporal reference, Computational Linguistics, 14:2, 15-28.
12. Van Lambalgen, M. and F. Hamm (2005). *The proper treatment of events*, Oxford: Blackwell.
13. Willems, R.M. et al. 2009. Neural dissociations between action verb understanding and motor imagery, Journal of Cognitive Neuroscience 22:10, 2387–2400.
14. Zwaan, R.A. and G.A. Radvansky (1998). Situation models in language comprehension and memory, Psychological Bulletin, 123:2, 162-185.