A Dynamic Update Model of Sentence Processing

Abstract. In recent years formal semantic theories based on the principle of compositionality have been criticized of not being able to explain how meaning is computed in the brain (Baggio and Hagoort 2011). For example, although neither *John squeezed an orange* nor *John squeezed an apple* contain a semantic anomaly, they are processed differently in the brain, because *orange* is more expected as the direct object of *squeeze* than *apple*. Alternative approaches, like that of Van Lambalgen and Hamm (2005) abandon compositionality by allowing lexical meanings to be both enriched and reduced during semantic processing.
In this article we try to bridge the gap between neuroscience and formal semantics. We develop a compositional dynamic update model of sentence processing that is based on strategies from Dynamic Epistemic Logic and Epistemic Temporal Logic. For example, the meaning of a verb is a set of evolutions, modeled by a branching tree that describes how events can evolve. The branches of the tree are ordered by a preorder that captures a comprehender's conditional beliefs about the plausibility of the different evolutions. New information is modeled as a (doxastic) event model. The result of updating the current discourse model with this new information is defined as a form of priority update.

Keywords: formal vs. cognitive semantics, compositionality, branching time logic, dynamic epistemic logic, combining systems, N400, P600, SAN

1 Formal semantic theories and the challenge from event-related potentials

1.1 Three event-related potentials components related to semantic processing

According to many, if not most, current formal semantic theories, common nouns like 'dog' or 'paper' are basically analyzed as sets of objects. For example, 'dog' is first translated as the lambda-term $\lambda x.dog(x)$, which, in a second step, is interpreted as a subset of the domain, or, more precisely, as a function from this domain to the set of truth values. Similarly, using a (neo-)Davidsonian approach, verbs are interpreted as sets of events, or, alternatively, as relations between n-tuples of objects and an event. For example, 'write' can be translated as the lambda-term $\lambda y \lambda x \lambda e(write(e) \wedge theme(e, y) \wedge actor(e, x))$.

In recent years, such approaches to defining the semantics of basic lexical items have been criticized from neuroscience. According to Baggio and Hagoort (2011), those theories are 'by design insensitive to differences between words of the same syntactic category denoting objects of the same type' (Baggio and Hagoort:1343). As a consequence, they are inappropriate as a theory of semantic processing in the brain. This criticism is based on empirical results from neurophysiological and neuroimaging phenomena like the N400, the P600 or the SAN (sustained anterior negativity)[1], which are components of event-related potentials (ERP's) whose amplitudes are modulated by semantic complexity.

N400. Consider the examples in (1) and (2).

(1) a. Jenny put the sweet in her mouth after the lesson.
b. Jenny put the sweet in her pocket after the lesson.

(2) Every morning John makes himself a glass of freshly squeezed juice. He keeps his refrigerator stocked with (oranges/apples/carrots).

A formal semantic analysis of the sentences in (1) differs only in the sort of object assigned to the locative argument of the verb *put*: *mouth* versus *pocket*. Yet, when this sentence is uttered in a context where Jenny leaves the classroom after a lesson, Hagoort and Brown (1994) found a difference in the N400 between *mouth* and *pocket*, showing that there is a difference during processing in the brain that needs to be accounted for by formal semantic theories.

Sentences like (2) were used by Federmeier and Kutas (1999) in an ERP experiment also targeting the N400. The authors found an increasing N400 effect with the ordering 'oranges' $<$ 'apples' $<$ 'carrots'. According to one interpretation of the N400, this effect is closely related to predicting upcoming words in a sentence which is based on semantic relations between words in the memory component of the brain. For example, in (2) both 'apple' and 'carrot' trigger a larger N400 compared to 'orange' because the former are semantically less related to an event of squeezing a fruit than the latter (Kutas and Federmeier 2000, Federmeier and Laszlo 2009).

P600. There are types of semantic information that do not have an impact on the N400 but which trigger effects later on in the time course of semantic processing. For example, Kuperberg et al. (2003) presented participants with three types of sentences with either no violation, a thematic violation or a pragmatic one.

(3)	a. For breakfast the boys would only bury ...	pragmatic violation
	b. For breakfast the eggs would only eat ...	thematic violation
	c. For breakfast the boys would only eat ...	no violation

Whereas N400 amplitudes were larger for the pragmatic violation compared to the sentence with no violation, there were no significant differences in the amplitudes for the sentences with no violation and the sentence with a thematic violation. This result shows that pragmatic information is available at the time of the N400, whereas this does not hold for information about how thematic roles are bound to arguments (Kuperberg et al. 2003, Federmeier and Kutas

[1] For details on these components, see Baggio et al. (2011) and Baggio et al. (2008).

2009). However, this thematic violation is detected by the brain at a later stage of semantic processing, eliciting an enhanced posterior positivity (P600).

Sustained anterior negativity. A third ERP component that is related to semantic processing is the sustained anterior negativity (SAN). This component occurs after the P600. Its role in semantic processing is illustrated by the examples in (4). When native Dutch speakers read sentences like (4) in an ERP experiment, a higher SAN was found for (4b) compared to (4a) for the verb 'morste' (spilled) at the end of the sentence.

(4) a. Het meisje was een brief aan het schrijven toen haar vriendin koffie op het tafelkleed morste.
'The girl was writing a letter when her friend spilled coffee on the tablecloth.

b. Het meisje was een brief aan het schrijven toen haar vriendin koffie op het papier morste.
'The girl was writing a letter when her friend spilled coffee on the paper.

A related, but independent challenge to those posed by ERP components is the so-called binding problem. Meaning-related information is represented in the brain in a highly distributed fashion in the sense that multiple brain areas are involved both in semantic access and semantic processing (Federmeier and Kutas 2009:3). For example, color and sound are represented separately from the instant they are perceived of an object, whereas color and shape are first encoded together and only subsequently analyzed by different areas in the brain (Holcombe 2009). This distribution of semantic information leads to the question of how such distributed feature information is brought together to create an integrated concept that comprehenders can use in processing spoken or written words like 'dog' or 'paper' (Federmeier and Laszlo 2009, Holcombe 2009). As emphasized in Hagoort (2005:418), this problem is even more intriguing for semantic information because one also has to account for the fact that this information is in addition processed at different time scales and at relatively widely spaced parts of the time axis. For example, the three ERP components discussed above occur at different times during semantic processing.

2 Two ways of solving the challenge from ERP components

2.1 A non-compositional account: Van Lambalgen and Hamm (2005)

Baggio and Hagoort (2011) argue that what is stored in memory are *unification-ready structures* like the scenarios that encode verb phrase meanings of Van Lambalgen and Hamm (2005). Such structures are stored in the memory component of the brain and are represented by a set of clauses (constraints), which together determine how the scenario can combine with other constructions to form a more complex construction.

Using this framework, the SAN effect for (4) is explained by Baggio, Van Lambalgen and Hagoort (2008) as follows. For the verb phrase *was writing a letter*, a (simplified) scenario consists of the statements in (5).

(5) a. $Initiates(start_{write}, write, t)$
b. $Terminates(finish_{write}, write, t)$
c. $HoldsAt(write, t) \wedge HoldsAt(letter(c), t) \rightarrow Happens(finish, t)$
d. $HoldsAt(write(x), t) \rightarrow Trajectory(write, t, write(x + g(d)), d)$
e. $HoldsAt(build, R), R < now$

The second and the third clause introduce a goal that determines a definite end state for the writing event. (5c) states that the writing activity comes to an end if the letter has reached a stage c where it is considered to be finished. The dynamic aspect of a writing event is captured by the second clause. If at t the theme of the writing exists to degree x, then the writing activity that lasts for interval d leads to the theme's increasing to degree $x + g(d)$, where g is a monotone increasing function relating the writing activity to the stage its theme is in. Clause (5e) is the contribution of the past progressive. By only considering the minimal model of the axioms of the event calculus plus the clauses of the scenario, one (non-monotonically) infers that the letter was eventually finished (Van Lambalgen and Hamm 2005). This inference is non-monotonic because it is possible to add a terminating event to the scenario and thereby cancel the inference that the letter was finished. Exactly this happens in the situation described by (4b) when the coffee is spilled on the paper. As a result, (5b) and (5c) (for $t = t'$) have to be retracted. Thus, the meaning of *write a letter* is changed 'outside-in, top down' due to new information further down the sentence. Since constraints can be added or retracted when two expressions are combined, the principle of compositionality does not hold.

2.2 A compositional account: a multi-layered framework for sentence processing

In the rest of the paper we will develop a compositional alternative that is able to cope with the data from section 1. Recall the major challenges posed by ERP components and the binding problem: (i) the meaning of lexical items cannot be reduced to sortal information plus information about the argument structure, i.e. information about the arity of the function that is used to interpret the item ('insensitive to differences between words of the same syntactic category'); (ii) integrated concepts are composed of features whose neurophysiological correlates are stored and processed in various brain areas and (iii) meaning is processed along the time axis in different, clearly separated stages.

The main tenets of our proposal are the following. First, the meaning of lexical items consists of several interrelated components. Each component is modeled by its own structure together with a language (or logic) to talk about this structure. Second, the different structures are related in various ways, allowing a flow of information between those structures. This leads to a multi-layered framework in the sense of the theory of combining system (Blackburn and de Rijke 1997). For example, concepts of persistent objects and events are represented as different structures, contributing sortal information. This accounts for the perspective of those entities as indivisible atoms. Entities as 'bundles' of features are defined as a different structure. Both sorts of structures are combined by assigning to each entity viewed as an indivisible atom a 'bundle'-structure determining, e.g. with respect to which attributes changes can be brought about (event) or be triggered by events (objects). Third, semantic processing is inherently *non-deterministic*. When processing a constituent, there are in general several options at the semantic level with what kinds of constituents it can combine with (difference to the syntactic level). Following dynamic semantics, one can say that the meaning of a lexical item is its context change potential (or a set of continuations). However, in contrast to current dynamic theories like Dynamic Predicate Logic or Update Semantics, the following points have to be stressed: (a) there are different dimensions at which lexical items can have a context change potential and (b) for each dimension, it is not simply a yes-or-no question which option is chosen but the options are ranked according to some order of, say, typicality. In order to account for those two constraints, lexical items are not interpreted as as objects in some global model, or as relations between states as in DPL, where a state is a function from a set X of variables to the universe M of some fixed first-order model $\mathcal{M}$. Rather, we follow Dynamic Epistemic Logic (Baltag, Van Ditmarsch and Moss 2008 for an overview) and interpret the meaning components of lexical items as a particular type of model, similar to the event (or action) models of Dynamic Epistemic Logic and the plusibility models of Baltag and Smets (2006, 2008). This move makes it possible to get a unified form of meaning across the different dimensions and rankings with respect to typicality inside one meaning component, like e.g. those between possible theme arguments of 'squeeze', are accounted for by defining orderings on the respective models with respect to which those meaning components are defined. This step makes it necessary to explicitly represent *update operations* in the formal model.

Consider the following two examples. After having processed *John squeezed* in *John squeezed an orange*, a comprehender knows that there are different continuations of the sentence with respect to the event in which he is involved. However, some types of events are more expected than others, say run or walk compared to jump and swim. This knowledge is independent of syntactic information ('John' is a proper name) and cannot be inferred solely on the basis of sortal information. At this stage of processing, she only has partial knowledge about the situation. As an effect, she is not able to determine a *unique* situation. Rather, there are different situations $s_1, ..., s_n$ that are possible according to the factual information she has got so far. A second source of non-determinism arises at the *dynamic* dimension of an event. Consider again the sentence *A girl was writing a letter*. After processing this sentence, a comprehender knows that (i) a writing event is described as ongoing, (ii) the actor is a girl and (iii) the theme is a letter. However, she does not yet know how this writing event evolves after the time of reference. At least the continuations in (6) are possible.

(6) possible continuations

a.	... After twenty minutes the letter was finished.	completed
b.	... when she had to stop because she had to give a lecture.	interrupted
c.	... when she died of a heart attack.	failed

The options a comprehender considers possible after having processed *The girl was writing a letter* cannot simply be deduced from the sortal information. Rather, they are directly related to the 'bundle'-view of the corresponding type of objects. In this case they have to do with the typicality or expectancy of how ongoing events of writing are (most) likely to evolve. For an accomplishment verb like 'write', a possible ordering is (7).

(7) accomplishment verb: completed $<$ ongoing $<$ interrupted $<$ failed

3 Formalizing multi-layered frameworks for sentence processing

3.1 Objects and events

Objects and events as atoms: capturing sortal information. Sortal information about (persistent) objects and events is captured by object and event structures, respectively. At this level, objects and events are taken as indivisible, atomic entities in the sense that they are not related to possible attributes (features) but only to other entities of the same type.

Definition 1 (Object structure) *An object structure $\mathcal{O}$ is a triple $\langle O, \{P_{cn} \mid cn \in CN, \sqsubseteq_O\}\rangle$ s.t. (i) O is a non-empty set of (persistent) objects, (ii) each P_{cn} is a subset of O with CN a set of object labels including 'dog' or 'paper' and (iii) $\sqsubseteq_O$ is the part-of relation on O.*

Definition 2 (Event structure) *An event structure $\mathcal{E}$ is a triple $\langle E, \{P_v \mid v \in Verb\}, \sqsubseteq_E\rangle$ s.t. (i) E is a non-empty set of (atomic) events; (ii) each P_v is a subset of E with $Verb$ a set of event labels including 'write' or 'squeeze' and (iii) $\sqsubseteq_E$ is the part-of relation on E.*

Scale transition structures. As already mentioned, both types of structures are used to express sortal information about indivisible atoms. They do *not* capture relations that exist between the two domains. For example, an event of squeezing brings about a change with respect to attributes or properties of the objects participating in it. What is in addition needed, therefore, are structures that model this mutual relationship between the two domains. This is done by *scale transition structures.*

Definition 3 (Scale transition structure) *A scale transition structure* $\mathcal{STS}$ *based on an attribute label set* Σ *is a triple* $\langle \mathcal{T}, \{E_\sigma \mid \sigma \in \Sigma\}, V\rangle$ *s.t. (i)* $\mathcal{T}$ *is a rooted tree (see Definition 6 below); (ii)* Σ *is a set of scale parameters like heat, height or length, (iii) each* E_σ *is a function from* T*, the domain of* $\mathcal{T}$*, to* $\wp(T)$ *and (iv)* V *is a valuation function that assigns to each node in* T *a boolean combination of elements from a set* $DEGREE$ *of propositional variables which are taken as degrees, i.e. the values corresponding to particular scales.*

The elements of Σ are the types of attributes or properties with respect to which an event can bring about a change during its occurrence like height or the stage an object is in while it is being created (say, a letter) or being destructed (say, an apple which is being eaten). Each E_σ models the 'powers' of an event e in the sense that if $s' \in X \in E_\sigma(s)$, then the state s' can be reached from the state s by an occurrence of an event e. If each $X \in E_\sigma(s)$ is linearly ordered, one gets a set of sequences (histories).

The relation between an event structure $\mathcal{E}$ and a scale transition structure $\mathcal{STS}$ is given by a function γ which assigns to each event $e \in E$ an $\mathcal{STS}$. In addition, it is required that events of the same type are assigned the same $\mathcal{STS}$: $\forall e \forall e' : e \in P_v \wedge e' \in P_v \rightarrow \gamma(e) = \gamma(e')$. One way of thinking about γ is to associate with each arc in the tree $\mathcal{T}$ a label v with the intuitive meaning that the transition between the values of the attribute σ is triggered by the event e of type v.

In addition to events, scale transition systems can also be used for the domain O of persistent objects. In this case, the scales represent attributes or properties together with the values such attributes can take. More importantly, the functions E_σ model the power of events that can change the values of the attributes. This relationship is defined by a function γ' which assigns to each element $o \in O$ of an object structure $\mathcal{O}$ a scale structure $\mathcal{STS}$.

So far, we deliberately left out the following two questions (i) with respect to what kinds of objects events of a given type (preferably) bring about a change and (ii) what kinds of events (preferably) effect changes in the attributes of an object. This kind of information is required to account for effects occurring with respect to both the N400 and the P600.

Definition 4 ($\mathcal{L}$-decorated scale transition structure) *An* $\mathcal{L}$*-decorated scale transition structure* $\mathcal{STS}_\mathcal{L}$ *is a scale transition structure whose arcs are labeled by elements of the label set* $\mathcal{L}$*. For a given* $\sigma \in \Sigma$*, the arcs are labeled by a subset* $\mathcal{L}'$ *of* $\mathcal{L}$*, denoted by* $\sigma(\mathcal{L})$*. Thus, an* $\mathcal{L}$*-decorated scale transition structure is a scale transition structure where instead of the set* $\{E_\sigma \mid \sigma \in \Sigma\}$*, one has the set* $\{E_{\sigma,cn} \mid \sigma \in \Sigma, cn \in \sigma(\mathcal{L})\}$*.*

For events, the label set $\mathcal{L}$ is CN, i.e. the labels for common nouns like 'dog'. Conversely, for objects one has $\mathcal{L} = Verb$. We illustrate the working of $\mathcal{L}$-decorated scale transition structure for events. The case for objects is analogous.

Let $\rightarrow_{\sigma,cn} = \{(t,t') \mid (t,t') \in E_{\sigma,cn}\}$ be the transition relation corresponding to $E_{\sigma,cn}$. For a $\sigma \in \Sigma$, one defines a (pre-)order $\preceq_{\sigma,\mathcal{L}}$ on the $\rightarrow_{\sigma,cn}$ with $cn \in$

$\sigma(\mathcal{CN})$ (this order must, of course, be determined empirically). Intuitively, if $\rightarrow_{\sigma,cn} \preceq_{\sigma,\mathcal{CN}} \rightarrow_{\sigma,cn'}$, this means that it is more plausible (expected) that an event e brings about a change with respect to the attribute σ if the object is of type cn than if the object is of type cn'. For example, for an event of type squeezing and σ = volume, part of the ordering is (using cn instead of $\rightarrow_{\sigma,cn}$ for better readability): $orange \preceq_{volume,\mathcal{CN}} apple \preceq_{volume,\mathcal{CN}} carrot$. For the common noun 'orange', one gets: $squeeze \preceq_{volume,\mathcal{VERB}} eat \preceq_{volume,\mathcal{VERB}} buy$.

Next we need define an ordering on CN-decorated $\mathcal{STS}'s$ for verbs and $VERB$-decorated $\mathcal{STS}'s$ for common nouns. Let $E_v(\Sigma)$ be the set of attributes with respect to which events of type v can bring about changes, i.e. it is that subset Σ' of Σ such that for some $e \in E$ of type $v \in VERB$, $\gamma(e) = \mathcal{STS}$ and $\mathcal{STS}$ is based on Σ'. Similarly, $O_{cn}(\Sigma)$ is the set of attributes defined for an object $o \in O$ of type cn. A limiting case in the combination of a common noun of type cn and a verb of type v is given by $E_v(\Sigma) \cap O_{cn}(\Sigma) = \emptyset$. In this case, a larger N400 is triggered (pragmatic violation). In the more general case, i.e. if $E_v(\Sigma) \cap O_{cn}(\Sigma) \neq \emptyset$, the amplitude of the N400 depends on the 'semantic fit' between the verb and the common noun. This 'fitness' relation is defined as follows.

Definition 5 (Semantic fit between verbs and common nouns) *For a given $\sigma \in \Sigma$, label sets $VERB$ and CN with $\sigma \in E_v(\Sigma) \cap O_{cn}(\Sigma)$, the ordering $\preceq_\sigma$ on $(\preceq_{\sigma,CN} \times \preceq_{\sigma,VERB})^2$ is defined by: $\langle \rightarrow_{cn}, \rightarrow_v \rangle \preceq_\sigma \langle \rightarrow_{cn'}, \rightarrow_{v'} \rangle$ iff $\rightarrow_{cn} \preceq_{\sigma,CN} \rightarrow_{cn'}$ and $\rightarrow_v \preceq_{\sigma,VERB} \rightarrow_{v'}$.*

According to Definition (5), the ordering $\preceq_\sigma$ depends both on $\preceq_{\sigma,VERB}$ and $\preceq_{\sigma,CN}$ in the sense that no parameter is preferred. Other orderings are possible as well. Examples are orderings based on lexicographic or anti-lexicographic order. Another alternative is to use the sum of the actual positions in the respective orderings: $\langle \rightarrow_{cn}, \rightarrow_v \rangle \preceq_\sigma \langle \rightarrow_{cn'}, \rightarrow_{v'} \rangle$ iff $(pos(\rightarrow_{cn}, \preceq_{\sigma,CN}) + pos(\rightarrow_v, \preceq_{\sigma,VERB}$ $)) \leq (pos(\rightarrow_{v'}, \preceq_{\sigma,VERB}) + pos(\rightarrow_{v'}, \preceq_{\sigma,VERB}))$. Since the exact ordering is a matter of empirical investigation, we will stick with the ordering given in Definition (5). Let us illustrate the working of $\preceq_\sigma$ by the example in (8).[2]

(8) a. squeeze: $orange \preceq_{volume,\mathcal{CN}} apple$

b. orange : $squeeze \preceq_{volume,\mathcal{VERB}} eat$

c. (squeeze,orange) : $(orange, squeeze) \preceq_{volume} \left\{ \begin{array}{c} (orange, eat) \\ (apple, squeeze) \end{array} \right\} \preceq_{volume}$ $(apple, eat)$

[2] This example shows that we need to combine different plausibility orderings. This raises the question whether it is better to use quantitative or even probabilistic rankings. Such a move would make it possible to also define an addition operation on different orderings. As a consequence, it becomes possible to define the update operation of the current state during semantic processing with the next upcoming item in terms of the sum of the probability of the expected continuations defined by the current state and the probability of that state defined by the upcoming item. We must leave this question to another occasion.

The lowest N400 is triggered if 'orange' is combined with 'squeeze' because it is the minimal element of $\preceq_\sigma$. The highest N400 is expected to occur with the pair $(apple, eat)$. More generally, let $R \subseteq \preceq_\sigma$. $Min_{\preceq_\sigma} = \{(cn, v) \mid (cn, v) \preceq_\sigma (cn', v')$ for all $(cn', v') \in \preceq_\sigma\}$ is the set of all minimal elements of $\preceq_\sigma$. Then the elements which semantically fit best are in this set. By contrast, for the P600, the question is *not* that of to what degree a verb and an argument semantically fit. What is tested, rather, is the condition that no single attribute σ is assigned to two different pairs (cn, v) for a given verb v.

The N400 is only sensitive to the ordering $\preceq_\sigma$: the better the semantic fit, i.e. the lower in the ordering a pair is found, the lower this effect. This does *not* exclude that there are different $cn's$ assigned to the same σ. If this happens to be the case, this has the effect of an attribute being assigned to two different objects. Consider again (3b) 'For breakfast, the eggs eat ...'. The problem in this case is the fact that a possible continuation like 'cornflakes' also semantically best fits with the 'volume' attribute, similar to 'eggs'. As a result, this attribute is assigned to two different arguments, a fact that is not allowed (cf. thematic uniqueness).

3.2 Event evolution trees

$\mathcal{STS}'s$ encode information about which results can be brought about by an event and which attributes of persistent objects can be changed by events. They do *not* encode *in what way* these results can be brought about and the different ways of how the occurrence of an event can be successful or not. For example, events can be described as *completed*, *ongoing* or *interrupted* (see example (4) from section 1). This information is encoded in *event evolution trees.*

Definition 6 (Tree) *A tree $\mathcal{T}$ is a pair $\langle T, < \rangle$, where T is a non-empty set of time points and $<$ is a strict partial ordering satisfying the tree condition: $\forall t \forall t_1 \forall t_2 (t_1 < t \wedge t_2 < t \rightarrow t_1 < t_2 \vee t_2 < t_1 \vee t_1 = t_2)$. A tree is rooted if it has a least element, called its root. The root will be denoted by r.*

Definition 7 (History and branch) *Let $\mathcal{T} = \langle T, < \rangle$ be a tree. A history in $\mathcal{T}$ is a maximal linearly ordered set of time points. A branch in $\mathcal{T}$ is any set of time-points $\{t | t \in h$ and $t' < t\}$ for some history h and time point t'. The least node of a branch b is denoted by $I(b)$ and called the initial time-point of b. When $\mathcal{T}$ is a rooted tree, every history is a (maximal) branch in $\mathcal{T}$. The history h is said to pass through t whenever h contains t, i.e. $t \in h$. Two branches are siblings if they have the same initial time-point. Siblinghood is an equivalence relation on the set of branches in a tree and is denoted by $\sim$.*

The set of all branches in a tree $\mathcal{T}$ is denoted by $B(\mathcal{T})$ and the set of all histories in a tree by $H(\mathcal{T})$.

Definition 8 (Event evolution tree) *An event evolution tree $E(T)$ is a quadruple $\langle \mathcal{T}, \mathcal{E}, \alpha, \beta \rangle$ where $\mathcal{T}$ is a rooted tree, $\mathcal{E}$ is an event structure, and α and β are total functions from E to T that assign to an event $e \in E$ its beginning point $\alpha(e)$ and its end point $\beta(e)$, respectively.*

In terms of α and β a transition relation $\lhd_e$ can be defined on $T \times T$ by setting $(t, t') \in \lhd_e$ iff $\exists e \in E$ $s.t.$ $\alpha(e) = t$ and $\beta(e) = t'$. One has: $\lhd_e \subseteq <$. Furthermore, $\lhd_e$ can be defined as the union of the sorted relations $\lhd_{e,v}$ with $v \in VERB$: $\bigcup_{v \in VERB} \lhd_{e,v} = \lhd_e$ so that each $\lhd_{e,v}$ is a subset of $\lhd_e$. Whenever one has $(t, t') \in \lhd_e$, this means that on the interval between t and t' an event of the given type, say a writing, occurred. However, there can be pairs (t, t') s.t. no event occurred during that time span.

For an event evolution tree $E(T)$, the set of histories $H(\mathcal{T})$ is ordered by a well-preorder $\preceq$, the *plausibility* or *expectancy* ordering on histories (evolutions). A well-preorder is a preorder, i.e. a reflexive and transitive relation, such that every non-empty subset has minimal elements. Thus, if

$$Min_{\preceq} R := \{t \in R \,|\, t \preceq t' \text{ for all } t' \in R\}$$

is the set of minimal elements of R, the condition on non-emptiness requires that: for every $R \subseteq H(\mathcal{T})$, if $R \neq \emptyset$ then $Min_{\preceq} R \neq \emptyset$. If $H(\mathcal{T})$ is finite, a well-preorder is simply a connected preorder.

The language $\mathcal{L}$ for talking about event evolution trees is recursively defined by $\phi = p | \neg\phi \,|\, \phi \wedge \phi \,|\, P\phi \,|\, F\phi \,|\, \Diamond\phi \,|\, \langle \lhd_e \rangle \phi$, with (i) $p \in DEGREE$ are a denumerable set of propositional degree variables that are used to make assertions about the value of an attribute like 'height' or the current state of a letter which is being written; (ii) F and P are the temporal modalities for future and past and (iii) the modality $\diamond$ (*possible alternative history*) and (iv) $\langle \lhd_e \rangle \phi$ (*occurrence of an (atomic) event*). As usual, $G, H, \Box, [\lhd_e]$ are defined as the duals of $F, P, \Diamond, \langle \lhd_e \rangle$; respectively. Formulas of $\mathcal{L}$ are evaluated on pairs $\langle h, t \rangle$, where t is a time-point and h is a history passing through t. The underlying intuition is that the truth of a formula depends not only on a time point but also on a particular history to which this time point belongs.

Definition 9 (Valuation on an event evolution tree) *A valuation on an event evolution tree $E(T)$ is any mapping V that assigns to each propositional variable in PV a set of pairs $\langle h, t \rangle$.*

A *model* is a pair $\langle E(T), V \rangle$ where $E(T)$ is an event evolution tree and V a valuation in $E(T)$. Truth of a formula in a model $\mathcal{M}$ relative to a pair $\langle h, t \rangle$ is defined as follows.

- $\mathcal{M}, h, t \models p$ iff $\langle h, t \rangle \in V(p)$
- $\mathcal{M}, h, t \models \neg\phi$ iff not $\mathcal{M}, h, t \models \phi$
- $\mathcal{M}, h, t \models \phi \wedge \psi$ iff $\mathcal{M}, h, t \models \phi$ and $\mathcal{M}, h, t \models \psi$
- $\mathcal{M}, h, t \models F\phi$ iff $\mathcal{M}, h, t' \models \phi$ for some t' with $t < t'$
- $\mathcal{M}, h, t \models P\phi$ iff $\mathcal{M}, h, t' \models \phi$ for some t' with $t > t'$
- $\mathcal{M}, h, t \models \Diamond\phi$ iff $\mathcal{M}, h', t \models \phi$ for some h' passing through t
- $\mathcal{M}, h, t \models \langle \lhd_e \rangle \phi$ iff $\mathcal{M}, h, t' \models \phi$ for some t' with $(t, t') \in \lhd_e$
- $\mathcal{M}, h, t \models \langle \lhd_{e,v} \rangle \phi$ iff $\mathcal{M}, h, t' \models \phi$ for some t' with $(t, t') \in \lhd_{e,v}$

In order to define different types of evolutions, the following program constructs from Propositional Dynamic Logic (PDL) are used for $\lhd_e$:[3] (i) sequential composition ; $[[\lhd;\lhd]] = [[\lhd_e]] \circ [[\lhd_e]]$; (ii) Kleene star $\lhd_e^*$, which is defined as the reflexive and transitive closure of $\lhd_e^n$ and $\lhd_e^+ = \lhd_e; \lhd_e^*$; (iii) $\cup$ (choice): $[[\lhd_e \cup \lhd_e]] = [[\lhd_e]] \cup [[\lhd_e]]$. The test program $?\phi$ is defined by $[[?\phi]] = \{(\langle h,t\rangle, \langle h,t\rangle) \mid \mathcal{M}, h, t \models \phi\}$ with ϕ a propositional variable.

Definition 10 (Evolution in an event evolution tree) *An evolution in an event evolution tree $E(T)$ of type π is a history h satisfying a particular type of formula at the root r, the characterizing formula for evolutions of type π.*

The characterizing formula is formulated in terms of $\lhd_e$ and the programming constructs from Propositional Dynamic Logic because they refer to the development of an event of the type described by the event evolution tree. For accomplishment and activity verbs, one gets the following types of evolutions, corresponding to (7a).

(9) a. completed: $\langle(?(\neg\phi); \lhd_e)^+; ?\phi\rangle\top$
b. ongoing: $\langle\lhd_e^+\rangle\Box\langle\lhd_e\rangle\top$
c. interrupted: $\langle\lhd_e^+\rangle\Box[\lhd_e]\bot$
d. failed: $\langle\lhd_e^+\rangle\Diamond\langle\lhd_e\rangle\top$

A completed evolution requires the goal or result ϕ to be brought about. For the three other types, there is no such requirement. However, in contrast to a completed evolution, they not only impose a condition on the current history but also on its siblings after the condition on the current history is satisfied. For an ongoing evolution, used in the progressive, it is possible to continue the event sequence at all siblings. By contrast, for a failing evolution, no continuation of the event is possible at all, i.e. this type of event is disabled at all siblings. Finally, for interrupted evolutions, there is at least one possible continuation. For all histories in the event evolution tree of an accomplishment or an activity verb, (10) holds, which states that any evolution of an event of type v consists of at least the onset and a development portion.

(10) $\mathcal{M}, h, r \models \Box\langle\lhd_e^+\rangle\top$

3.3 Plausibility models

Plausibility models are used both for representing the meaning of lexical items and (ii) the operations that are used to combine those items to larger units (updating operations). This is achieved first by combining the different structures that have been defined in the preceding sections for representing (parts of) the meanig of lexical items, and second, by defining appropriate updating operations.

We start by giving the definition of a plausibility model (see Baltag and Smets 2006, 2008, Dégremont 2010).

[3] Note that $\lhd_e$ is used both for the modality and the corresponding relation.

Definition 11 (Plausibility model) *A plausibility model based on a set Prop of propositional letters is a triple* $\mathcal{M} = \langle W, \leq, V \rangle$ *s.t. (i)* $W \neq \emptyset$*, (ii)* $\leq$ *is a well-preorder on* W *and (iii)* $V : Prop \rightarrow \wp(W)$.

One has for the initial plausibility model a comprehender starts with when constructing a model of the situation described by a sentence, one has: (i) $W = \{w_0\}$; (ii) $\leq = \{(\lhd_e^+, \lhd_e^+)\}$ and (iii) the valuation V is determined by the context. When the comprehender successively parses the constituents of the sentence, this initial model is repeatedly updated with the semantic information provided by those constituents. As a result, one gets a set of sequences of models that represents the steps a comprehender can take when constructing a discourse model of the situation described. The semantic information provided by lexical items is itself represented by plausibility models. Since there are (possibly) different meaning components, the semantic contribution of a lexical item will in general correspond to different plausibility models (distributive character aus Section 1 aufnehmen). For example, a verb introduces (a) sortal information, it is a writing or a squeezing), (b) information about what kinds of changes can be brought about with respect to which attributes together with a plausibility ordering on them (STS) and (c) information about possible evolutions and a plausibility ordering on them (EET).

'Events' in a plausibility event model represent part of the meaning of a linguistic item and not real events. Despite this incongruence, we call them 'events' because for a comprehender each part or constituent of a spoken or written sentence is like an event of a public announcement in the sense of Dynamic Epistemic Logic that triggers an update in her current information state.

Each state w in a plausibility model corresponds to a particular way of how one type of meaning component can be interpreted. If a model has more than one element, this represents a plausibility ordering on possible continuations. The elements ('events') of W are the characterizing formulas of the respective structures corresponding to the various meaning components. These relations are captured by functions that assign to each state of a plausibility model a characterizing formula. For example, in the plausibility model defining the dynamic component of an event corresponding to an EET, this relation between a plausibility model and an event evolution model is captured by a function ρ which assigns to each state w in a plausibility model the characterizing formula of an evolution. For example, for an accomplishment verb like *write*, a plausibility model has four states $w_1, ..., w_4$ and, using the abbreviations from (9), one has $\rho(w_1) = completed, \rho(w_2) = ongoing, \rho(w_3) = interrupted, \rho(w_4) = failed$. For a verb, three types of plausibility models must be distinguished.

Definition 12 (Plausibility model for a verb: sortal information) *A plausibility model of a verb for sortal information is a triple* $\langle \{p_v\}, \leq, pre \rangle$ *with* p_v *the sortal information and* $\leq = \{(p_v, p_v)\}$*. The precondition pre depends on the type of verb. For example, for write it requires that the object written not be finished. For dry the degree of dryness must not be zero.*

Definition 13 (Plausibility model for a verb: evolution information) *An event plausibility model for an in-adverbial is a triple* $\langle\{w_1, w_2, w_3, w_4\}\rangle$ *with* $w_1 = (?(\neg\phi); \lhd_e)^+; ?\phi\rangle\top\}$, $w_2 = \langle\lhd_e^+\rangle\Box\langle\lhd_e\rangle\top$, $w_3 = \langle\lhd_e^+\rangle\Box[\lhd_e]\bot$ *and* $w_4 = \langle\lhd_e^+\rangle\Diamond\langle\lhd_e\rangle\top$. *The ordering is given by* $w_1 \leq w_2 \leq w_3 \leq w_4$. *There is no precondition since this has already been defined by the model for sortal information.*

Similar to the characterizing formulas of plausibility models for evolution information, the characterizing formulas for plausibility models stating static defeasible information are defined in terms of $\lhd_e$ and its subrelations $\lhd_{e,v}$. Note first that each $\lhd_{e,v}$ transition occurs with respect to at least one $\sigma \in E_v(\Sigma)$. Since $E_v(\Sigma)$ is finite, it is possible to define transitions $\lhd_{e,v,\sigma}$ s.t. (i) $\lhd_{e,v,\sigma} \subseteq \lhd_{e,v}$ and (ii) $\bigcup_{\sigma\in E_v(\Sigma)}$. The characterizing formulas are of the form $cn \to [\lhd_{e,v,\sigma}]cn$, expressing that cn is an invariant of $\lhd_{e,v,\sigma}$ transitions.

Definition 14 (Plausibility model for a verb: static defeasible information for attribute σ) *A plausibility model of a verb for static defeasible information for an attribute* σ *is a triple* $\langle W\rangle$ *with (i)* $W = \bigcup_{cn\in E_v(\Sigma)} cn \to [\lhd_{e,v,\sigma}]cn,$*, (ii)* $\leq$ *is induced by* $\preceq_{\sigma,CN}$*: if* $\to_{\sigma,cn} \preceq_{\sigma,\mathcal{CN}} \to_{\sigma,cn'}$*, then* $cn \to [\lhd_{e,v,\sigma}]cn \leq cn' \to [\lhd_{e,v,\sigma}]cn'$ *and (iii) the preconditions are those of the plausibility model for sortal information.*

The plausibility model for a *for*-adverbial differs from that of a verb for its evolution information only in the ordering. Instead of *completed* $\leq$ *ongoing* $\leq$ *interrupted* $\leq$ $failed$, one has *ongoing* $\leq$ *interrupted* $\leq$ $failed$ $\leq$ *completed.*

The updating operation. Updating a plausibility model with a plausibility event model is defined as *priority update* (see Dégremont (2010) and Baltag and Smets (2006, 2008) for details).

Definition 15 (Priority update) *Priority update of a plausibility model* $\mathcal{M} = \langle W, \leq, V\rangle$ *relative to a protocol* ρ *and a plausibility event model* $\mathcal{PE} = \langle E, \leq, pre\rangle$ *is the plausibility model* $M \otimes \mathcal{PE} = \langle W', \leq', V'\rangle$ *defined by (i)* $W' = \{(w, e) \in W \times E \,|\, \mathcal{M}, w \models pre(e) \wedge e \in \rho(w)\}$*, (ii)* $(w, e) \leq' (w', e')$ *iff either* $e \leq e'$ *or* $e \cong e'$ *and* $w \leq w'$ *and (iii)* $V'((p, e)) = V(p)$.

As noted by Baltag and Smets (2006, 2008), the update rule defined in Definition 15 (ii) corresponds to *antilexicographic* update because the preferences of the plausibility event model have precedence over that of the model to which it is applied. In case the plausibility event model does not give any priority, the comprehender sticks with the prior preferences.

This construction is defined below. It is an adapted variant of the corresponding construction for epistemic models in Van Benthem et al. (2010). In Definition 16, $\sigma_{(k)}$ is the initial segment of σ of length k and σ_k is the kth component of σ. Furthermore, σ^L is a sequence of plausibility event models of length n for length$(\sigma) = n$ and σ^R is the corresponding sequence $e_1, ..., e_n$ of events.

Definition 16 (σ^L-generated model) *Let* $L = \langle W, \leq, V\rangle$ *be a plausibility model and* ρ *be a function assigning to each* $w \in W$ *a state-dependent protocol.*

Given a sequence σ of plausibility event models, the σ^L-generated model under ρ, $\mathcal{M}^{\sigma^L,\rho} = \langle W^{\sigma^L,\rho}, \leq^{\sigma^L,\rho}, V^{\sigma^L,\rho}\rangle$ is inductively defined on the initial segment σ^L: (i) $W^{\sigma^L_{(0)},\rho} = W$, $\leq^{\sigma^L_{(0)},\rho} = \leq$ and $V^{\sigma^L_{(0)},\rho} = V$, (ii) $w\tau \in W^{\sigma^L_{(n+1)},\rho}$ iff (1) $w \in W$, (2) $\sigma^L_{(n+1)} = \tau^L$, (3) $w\tau_{(n)} \in W^{\sigma^L_{(n)},\rho}$, (4) $\tau \in \rho(w)$ and (5) $\mathcal{M}^{\sigma^L_{(n)},\rho}, w\tau_{(n)} \models pre_{\tau^R_n}(\tau^L_{n+1})$, (iii) for each $w\tau, v\tau' \in W^{\sigma^L_{(n+1)},\rho}(0 < n < len(\sigma^L))$, $w\tau \leq^{\sigma(n+1)}$ iff either (i) $\tau^R_{n+1} \leq^{\sigma^{L,\rho}_{n+1}} (\tau^R_{n+1})'$ or $\tau^R_{n+1} \cong^{\sigma^{L,\rho}_{n+1}} (\tau_{n+1}R)'$ and $w\sigma_{(n)} \leq^{\sigma^{L,\rho}_{(n)}} v\sigma'_{(n)}$ and (iv) for each $p \in AT(L)$, $V^{\sigma^L_{(n+1)},\rho}(p) = \{w\sigma \in W^{\sigma^L_{(n+1)},\rho} \mid w \in V(p)\}$.

Definition 17 (Event generated plausibility model) *The model generated by Definition 16 $\mathcal{M} = \langle H, \leq', V'\rangle$ is defined as follows: (i) $H = \{h \mid$ there is a $w \in W, \sigma = \bigcup_{w\in W} \rho(w)$ with $= w\sigma \in W^{\sigma^{L,\rho}}\}$, (ii) for all $h, h' \in H$ with $h = w\sigma$ and $h' = v\sigma'$, $h \leq h'$ iff $w\sigma \leq^{\sigma^{L,\rho}} v\sigma'$ and (iii) for each $p \in At(\mathcal{L})$ and $h = w\sigma \in H$, $h \in V'(p)$ iff $h \in V^{\sigma^{L,\rho}}$.*

An event generated plausibility model branches if the domain of the updating $\mathcal{PE}$ has more than one element and therefore allows non-determinism. These models function as unification-ready structures because a comprehender can use them by plugging in the current information in form of characterizing formulas. In general, this need only be done for the sortal information because the other information is fixed by the relation between this type of information and the other meaning components. So far, Definition 17 does not involve a revision strategy that determines how a sequence is changed if a choice that has been made, leading to a particular transition considered to be most plausible, has to be changed if the next upcoming item does not satisfy this expection. So the question is: how can Definition 5 be integrated into the algorithm?. Directly related to this lacuna is a second one. Using the antilexicographic update, the item processed last determines the new ordering. However, this order independence need not always hold. E.g., a 'for'-adverbial imposes its ordering independently of the position in a sentence it occurs.

A possible language $\mathcal{L}$ for talking about *event generated plausibility models* contains the modalities A, B and $\Box$. The atomic formulas of $\mathcal{L}$ are the characterizing formulas of PDL assigned to the elements of W by the function ρ. Formulas of $\mathcal{L}$ are evaluated on histories $w\sigma$ with σ sequences of 'events' corresponding to the information provided by (dynamic) constituents. A and B are used to talk about the 'vertical' dimension, i.e. about histories belonging to the same depth of an event generated plausibility model. $\Box$ is used to talk about all extensions of a history and therefore about the 'horizontal' dimension. Let $w\sigma \sqsubseteq w\sigma'$ be true iff σ is a prefix of σ'. The clauses for A, B and $\Box$ are given in (11), where $\mathcal{M}$ is an event generated plausibilty model and $last(w\sigma)$ is the last element of $w\sigma$.

(11) a. $\mathcal{M}, w\sigma \models p$ iff p is a characterizing formula and $\mathcal{M}, last(v\sigma) \models p$

b. $\mathcal{M}, w\sigma \models A\phi$ iff $\mathcal{M}, v\sigma \models \phi$ for all $v\sigma \in Domain(\mathcal{M})$

c. $\mathcal{M}, w\sigma \models B\phi$ iff $\mathcal{M}, v\sigma \models \phi$ for all $v\sigma \in Domain(\mathcal{M})$ s.t. $v\sigma \in Min_{\leq}(w\sigma)$
d. $\mathcal{M}, w\sigma \models \Box\phi$ iff $\mathcal{M}, w\sigma' \models \phi$ for all σ' with $\sigma \sqsubseteq \sigma'$

The truth of a formula $A\Box\phi$ is independent of the processing stage if it holds at one state in the initial model, i.e. if one has $\mathcal{M}, w \models A\Box\phi$ for $w \in W$. By contrast, according to (11a) a sentence ϕ is believed at the n-th stage ($A\phi$), with n the length of $w\sigma$, if ϕ holds at all sequences of the same length which are minimal relative to the plausibility ordering $\leq$. The truth of such a sentence depends in general on the depth of the model.

How does an event generated plausibility model represent the actual processing of sentences? Sortal information triggers the activation of orderings like those in (8a,b). If processing a sentence has reached a state s, this type of ordering is used to pre-activate the most plausible transition (continuation). Thus, its function is *prospective*. By contrast, orderings based on Definition 5, illustrated in (8c), are used as tests: is the (defeasible) decision taken based on the first ordering corroborated by the actual continuation? The lower a matching pair is found in the ordering, the better the semantic 'fit' and the lower the amplitude of the N400. Thus, this type of ordering is used locally during the actual upating step involving two adjacent constituents (corresponding to single transitions in the model) because it is based on a combination of two type (i) orderings (N400 and $\mathcal{STS}$ or relations between $\mathcal{STS}'s$). Further down the sentence, more global constraints apply (P600, SAN), since decisions made higher up in the model are at stake. Similar to the N400, the P600 corresponds to $\mathcal{STS}'s$ and relations between $\mathcal{STS}'s$. It differs from the N400 with respect to the kind of relation that is used. N400: place in the ordering $\preceq_\sigma$ vs. P600: uniqueness of arguments assigned to an attribute that can be changed by the event. Thus, a P600 is triggered whenever two different objects are assigned to the same property. Defeasible decisions about possible evolutions of the event are tested by the corresponding plausibility models. Such decisions have to be revised whenever both the information provided by the verb and that of another constituent, e.g. a 'for'-adverbial in case of an accomplishment verb, are received.

3.4 Information change, non-monotonicity and compositionality

In our approach, there are two types of information a comprehender can get from an event generated plausibility model: (i) information that holds during each stage of the construction process. This information is safe in the sense that it will not be revised when further, additional true information is got when processing the sentence. An example is sortal information that is provided by event or object models; (ii) information that holds for the most plausible (expected) evolutions but not necessarily for all evolutions. This information is not safe because it can be overridden by other (true) information in a sentence. An example is the information that the goal will be attained for an accomplishment verb like *write*. However, this information is only satisfied for the most plausible evolution and can be overridden by *for*-adverbials and/or subordinate clauses like *but she didn't* *finish the letter*. Information of the second type can lead to *non-monotonicity*. Two principle cases have to be distinguished: (a) *discarding a type of evolution*. This corresponds to a 'hard' fact which excludes all alternative evolutions which do not satisfy the constraint imposed by the new information. An example is an *in*-adverbial as in *The girl wrote the letter in 20 minutes*, which excludes all evolutions in which the letter was not completed. (b) *change of the plausibility ordering*. In this case the new information is 'soft' in the sense that it does not exclude any of the evolutions envisaged so far, but it makes evolutions the most expected ones that were considered less plausible before. Consider again the case of a writing event. If all the comprehender knows is that it is a writing (initial model $\mathcal{M}$), she will assume that a completed writing will be described using her plausibility ordering on the event evolution tree for *write*. This is still the case after having processed *The girl wrote the letter*. If next the temporal adverbial *for twenty minutes* is encountered, the expectation about the completion has to be changed. Now it is more plausible that the culmination (completed letter) was not attained. As a consequence, she will change the ordering on the set of possible evolutions: each evolution in which the culmination is not attained becomes more plausible than the completed evolution where the letter is finished. This does *not* exclude that the culmination was not reached: *The girl wrote the letter for twenty minutes. Then she stopped because the letter was finished.* This example also shows that non-monotonicity does not imply a violation of compositionality. Non-monotonicity is the effect of having to choose at a particular point in the event generated plausibility model a transition which is *not* minimal in the current ordering or to have to go back to a point in the construction and choose a different, less expected transition. The overall model remains the same, though it is used by the comprehender in a different way.

4 Conclusion

In this article we presented a dynamic update model of sentence processing. The meaning of lexical items is defined in terms of plausibility models which encode expectancy relations of upcoming or already parsed constituents. A lexical item can be assigned more than one such model, reflecting the fact that semantic processing in the brain is a time-dependent, distributed process involving several separated stages. Instead of using operations like unification or functional application / composition to combine constituents, an explicit update operation is defined in the model. Preference relations test the semantic 'fit', which is a measure for the cost of integrating an upcoming expression into the current model. In our theory, this reflects the fact that preference relations are 'precompiled' so that having to take another route in the generated plausibility model triggers a recomputation. From the perspective of our theory, the failure of other approaches like that of van Lambalgen and Hamm to account for a compositional analysis is due to the fact that by restricting meanings to basically sortal information or by not taking into account relations between items in the lexicon, neither the distributive character nor the non-determinism inherent in semantic processing or composition is taken into consideration. Let me close by indicating three important directions for future work: (i) is it possible to classify plausibility models? For evolution information, one can use an aspectual classification; (ii) How can other lexical items like quantifiers, adjectives and determiners be integrated into the framework?; (iii) most important is the question of how to incorporate the 'revision' operations that are triggered if the semantic 'fit' is not optimal or a prior defeasible choice must be changed.

References

1. Baggio, G., M. van Lambalgen and P. Hagoort (2008). Computing and recomputing discourse models: An ERP study, *Journal of Memory and Language*, 22, 2131-2140.
2. Baggio, B. and P. Hagoort. The balance between memory and unification in semantics: A dynamic account of the N400. *Language and Cognitive Processes*, 26, 2011: 1338-1367.
3. Baltag, A. and S. Smets (2006). Dynamic belief revision over multi-agent plausibility models, In G. Bonanno et al. (Eds.), *Logic and the Foundations of Game and Decision Theory: Proceedings of LOFT06*, TLG, Amsterdam UP, 11-24.
4. Baltag, A. and S. Smets (2008). A qualitative theory of dynamic interactive belief revision, In G. Bonanno et al. (Eds.) *Logic and the Foundation of Game and Decision Theory: Proceedings LOFT07*, TLG, Amsterdam UP, 13-60.
5. Baltag A., H.P. van Ditmarsch, and L.S. Moss, Epistemic logic and information update, In P. Adriaans and J.F.A.K. van Benthem (Eds.), *Handbook on the Philosophy of Information* , Elsevier Science Publishers, 2008, 361-456.
6. Blackburn, P. and M. de Rijke (1997). Zooming in, zooming out, *Journal of Logic, Language, and Information*, 6, 5–31.
7. Dégremont, C. (2010). *The temporal mind: observations on the logics of belief change in interactive systems*, PhD thesis, ILLC, University of Amsterdam.
8. Federmeier, K.D. and M. Kutas (1999). A rose by another name: Long-term memory structure and sentence processing, *J. Mem. Lang.*, 41, 469-495.
9. Federmeier, K.D. and S. Laszlo (2009). Time for meaning: electrophysiology provides insights into the dynamics of representation and processing in semantic memory, In B. Ross (Ed.), *The Psychology of Learning and Motivation*, Vol. 51, Burlington: Academic Press, pp. 1-44.
10. Hagoort, P. (2003). How the brain solves the binding problem for language: a neurocomputational model of syntactic processing, *NeuroImage*, 20, S18-S29.
11. Hagoort, P. (2005) On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences*, 9, 416-423.
12. Hagoort, P. and C. Brown (1994). Brain responses to lexical ambiguity resolution and parsing, In C. Clifton et al. (Eds.), *Perspectives on sentence processing*, Lawrence Earlbaum, Hillsdale NJ, 45-81.
13. Holcombe, A.O. (2009). The binding problem, In E. Bruce Goldstein (Ed.), The Sage Encyclopedia of Perception.
14. Kuperberg, G.R., T. Sitnikova, D. Caplan, and P.J. Holcomb (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences, *Cognitive Brain Research*, 17(1), 117-129.
15. Van Benthem, J., J. Gerbrandy, T. Hoshi, and E. Pacuit (2010). Merging frameworks for interaction, *Journal of Philosophical Logic*, 38(5), 491-526.
16. Van Lambalgen, M. and F. Hamm (2005). *The proper Treatment of Events*. Oxford, Blackwell.