Boban M. Arsenijević[1]
University of Niš
Faculty of Philosophy, Serbian Language Department

УДК 81'42
Оригинални научни рад
Примљен: 2. 9. 2012.

A MULTIHIERARCHICAL DISCOURSE REPRESENTATION[2]

Abstract: The paper reports about some basic aspects of a broader research that is still in progress. Arguments are presented in favor of a multihierarchical formalism used in the representation of the linguistic discourse, compared to the rather standard linear models. The formalism takes a deep decompositional approach to semantics, assuming that it bottoms at the level of atomic concepts, but where this level is flexible, and can be redefined if the empirical base requires a different, finer, level of granularity. Having presented a minimal definition of the formalism, the paper briefly examines a selection of its relevant syntactic and semantic aspects.

Keywords: discourse representation, linearity, multihierarchical representation, network model, discourse update, discourse-information retrieval

Introduction

Although meanings of linguistic expressions all involve parts which are invariant and independent of any context, a natural language sentence receives its full interpretation only by incorporation into the relevant discourse. In this paper I discuss some properties of the discourse in the process of integration, especially whether the discourse is structured linearly or in some other way.

The discourse representation is taken as the representation of the relevant body of information, or more precisely of all the information in a certain communication situation that is contributed by explicit communication (i.e. by linguistic utterances), correctly assumed by all the interlocutors to be presupposed, or correctly seen as obvious from the immediate context of the communication situation. Hence, a discrete view of communication is taken, in which there are static states of the discourse between any two contributions (discourse updates), irrespective of whether these are made through uttering new sentences or perceiving changes in the direct context. While discourse updates come in a linear sequence, a discourse representation is static, and not necessarily linear.

Still, the standard view of the discourse representation among syntactitians and semanticists is based on the file-card metaphor (Heim 1982). The discourse is seen as a set of propositions organized by file-cards. Each file-card contains one proposition about one discourse referent, and file-cards about the same referent are grouped together. In this paper, I criticize this view and propose a better model of discourse-representation, which is less influenced by syntax and phonology and more concerned with the role of the discourse in the process of communication.

[1] Ćirila i Metodija 2, 18000 Niš, Srbija; b.arsenijevic@gmail.com.

[2] This research is part of the research projects FFI2010-15006 by MICINN, and OI178014 by MPNRS.

In section 2, the file-card metaphor view of discourse is criticized, leading to a proposal of a better discourse representation in section 3. Section 4 discusses some interesting properties of the proposed discourse representation, while section 5 targets the question of how many discourse representations per speaker may take part in a single communication situation, and whether they can combine into more complex structures. Section 6 concludes.

The file-card metaphor

Two main problems of the file-card metaphor that we target concern 1) its contamination by genuinely phonological structural properties and 2) its restriction to a direct update of only one discourse-referent. The file-card metaphor is strongly based on the topic-comment relation, which is typical of linguistic expressions and argued to have phonological origin (Carstairs-McCarthy 1999). Moreover, discourse representation crucially involves memory in which information is stored, and on which grammatical and other modules operate to produce and perceive linguistic expressions. The structure of the discourse in the file-card metaphor, basically a list of propositions, however, involves a strong component of linearity. While it is a natural property of the phonetic realization of language, and perhaps also of the grammatical computation, both clearly mapping onto the discourse update – linearity in the memory is less expected. The question emerges whether linearity really is a feature of our discourse representations, possibly as an influence of the grammar, or is it only a feature of our models of discourse representation, as an influence of our models of grammar?

The second problem of the file-card metaphor is that a sentence like (1) can directly update only one discourse referent: John, Mary or Mary's husband, and the other two only secondarily, if at all. In a non-linear discourse representation, an online update of all the discourse-referents would be possible, hence improving, and making more realistic, the capacities of the discourse update.

(1) John kissed Mary in front of her husband.

The general point is that the file-card metaphor is a result of reasoning about the discourse in the immediate vicinity of grammatical structures. The obvious question arising is whether the discourse is necessarily structurally so similar to language, or does it perhaps have an own structure, serving the special requirements that it faces, which are not identical to those faced by grammar. In this paper, we investigate the latter option, that the discourse is structured different than the grammar. This direction imposes a somewhat speculative methodology, less strongly based on the empirical data, but it is nevertheless an important direction to go, if only to question some views about the discourse that have been taken for granted. Our goal is, however, more ambitious: we are trying to propose a discourse representation that is less based on linguistic grounds, but then more directly connected with other cognitive domains such as spatial cognition, computation of the contextual information, memory retrieval etc.

The proposal

We propose a web-like model of discourse representation in which all referents are represented through sets of relations with other referents, and which will be referred to as the Multihierarchical Associative Discourse Representation (MADR). MADR builds on earlier proposals of network-shaped discourse representations, such as the Two-Level Semantics and Semantic Networks (see e.g. Bierwisch 1983, Herweg and Maienborn 1992), but takes a radical decompositional approach to natural language semantics. It involves three types: the basic type of properties, the basic type of relations and the derived type of referents. A referent is a set of properties. The set of properties is potentialy infinite, but in actuality finite in every discourse representation. Properties correspond to primitive, indecomposable concepts (e.g. [warm], [place], [pleasant]), and on a deeper level are assumed to correspond to patterns of neural activations. The set of relations involves only one relation, which itself bears no meaning, and is symmetric, intransitive and non-reflexive. We call it *association*. Every property instantiated in a discourse representation has to be related by this association to at least one other property in the same discourse representation. This other property cannot be the same property of the same referent, but can be the same property of another referent, or another property of the same referent. This gives a web-like discourse representation in which each referent is represented through a set of properties, which usually relate it to other referents in the discourse. A rough formalization is given in (2).

(2) Discourse = {R | R={Prop | Assoc (Prop, Prop') & Prop' ∈ R'}}

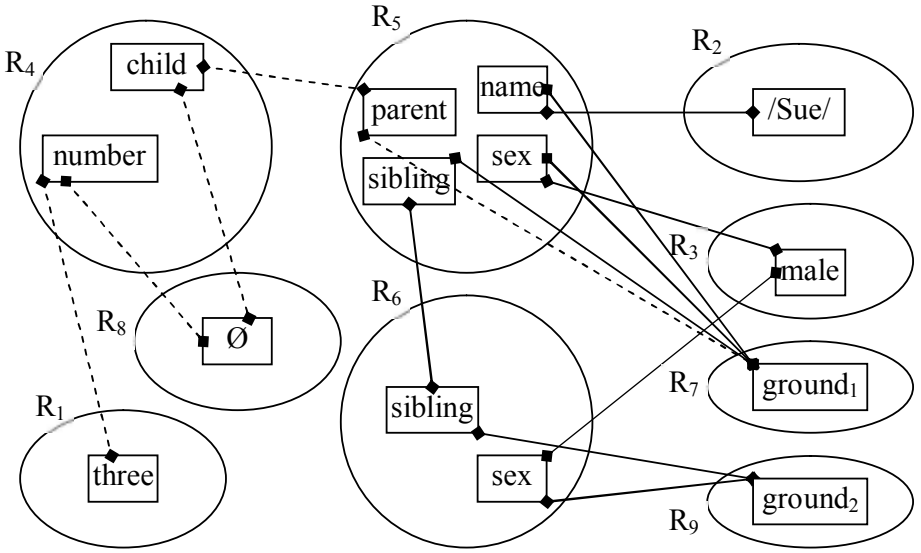
From the aspect of a particular referent, the relation between one of its properties and the property that it is associated with corresponds to the feature-value relation in grammar. The symmetry of the association implies that from the aspect of the other property and its referent, the feature-value relation inverts. In other words, the feature-value asymmetry is only introduced once we decide to look at some association from the angle of one of the associated properties. This symmetry of the association, with its other aspects, is a very important property of the proposed discourse representation and is discussed as the *multihiearchality* in more detail in section 4.

Discourse integration of new input retrieves the representation and either establishes new properties and associations or introduces new referents, while in production the retrieval targets certain discourse-prominent referents and relations and sends the retrieved material to grammar. We assume binary-branching grammatical structures in which each minimal meaning (contents of a head) takes two arguments (Dowty 1991 and van Valin & LaPolla 1997 for two-place predicate semantics). Every two-place predicate from the linguistic input results in a simultaneous update of two referents.

To illustrate this so far quite abstract description of MADR, let me construct a very simple discourse representation, in which there are nine referents: R_1, to R_9. Six of them are generic, i.e. without discourse-functions (R_1, R_2, R_3, R_7, R_8 and R_9) and three are individuals, with discourse-functions (R_4, R_5 and R_6). Instead of listing the specification of each of the referents, which would involve a lot of redundancy (because every defining associated property of every referent invokes another referent in which it has to be listed again), we present it graphically, as a web of referents organized in line with the formula in (2).[3]

[3] The contrast in economy between the algebraic, linear, representation of MADR, and the graphical

(3) MADR of a simple example of a discourse representation



Each of the rectangle boxes presents a property, each straight line with diamonds at its ends presents an association and each circle/ellipsis marks a set of properties, i.e. a referent. R_4 represents a divided referent ([number]) with the cardinality three, specified as children of R_5 ([child:parent]); R_5 represents a male referent named Sue, who is the father of R_5; R_6 represents a male referent, Sue's brother; R_4 is new to the discourse, i.e. it is being introduced at the point in time represented in the figure; R_6 and R_5 are part of the ground information (apart from the *parent* property, linked by the dotted line, which represents the absence of the relevant association in the shared discourse representation, and presence only in the interlocutor's own discourse representation, from which it is imported).

Properties associated with the same discourse role form an individual referent, and every property may form a singleton set representing a generic referent. This means that the discourse must dispose with as many discourse roles for its referents as the maximal number of individual, non-generic referents that it may take. However, discourse roles are assigned locally, in relatively small discourse domains, so that their number remains reasonably low. At discourse update, a new property, by which the representation of an individual referent is to be updated, is introduced by association with the respective discourse role.

Singleton sets represent basic generic notions; they are not associated with discourse roles, and even discourse roles themselves are represented as singleton sets. Every singleton set is a referent. The distinctive property of a generic referent is that it has no discourse function. This enables the definition of an additional class:

one, illustrates an important point. If segments of MADR indeed come close to the actual memory representations of sentential meanings, their straightforward linearization into sentences would lead to extreme redundancy and complexity. The mapping between MADR and a linear representation, done by syntax, must involve a lot of compression, mediated by the lexicon and the structure- and context-licensed deletion of certain material. This aspect is not discussed in the present paper.

complex generic referents, i.e. non-singleton generic referents. An instance of such a generic referent in the given example would be a set containing the properties *male* and *parent*, and representing the concept of a father. But what encodes the fact that certain properties form a set, and hence a complex generic referent (recall that for individual referents this is done by their discourse roles)? Two answers can be given to this question. One is that there is no marking of such sets, and that they are simply freely formed in any discourse. The problem of this answer is that it allows for the formation of many barely thinkable generic concepts, by grouping properties that are unlikely to represent a generic concept that is remotely part of our experience. This aspect may be regulated by other mechanisms, beyond the scope of this paper. If we still want to regulate this issue internal to the system, we may allow that complex generic concepts take only latent roles, or just roles from the domain of the long-term memory (instead of the discourse-roles), which allows them to be marked as sets in the long-term memory without receiving proper discourse roles in any particular discourse.

It may happen, however, that a generic referent is talked about (*I like red, it is so exciting*). In such cases, it must be assigned a proper discourse role. For this purpose, a new individual referent has to be formed by associating all the members of the generic referent with a discourse role. Discourse roles are treated as generic meanings with a special effect: they take part in specifying an individual discourse referent. An individual discourse referent is formed when a number of properties are associated with a single discourse function. Referents that are part of the interlocutor's individual knowledge, but not part of the shared discourse, have latent roles (for simplicity, the only such role in the example in (3) is marked as Ø), which define them as individual referents. Latent roles turn into proper discourse roles when the referent is introduced into the discourse from the interlocutor's individual knowledge.

The example in (3) presents a very poor discourse, with a small number of referents and without, for instance, any eventuality, or any temporal or spatial specification. This is done deliberately, for reasons of simplicity and clarity (which are already with such a simple discourse significantly threatened). The addition of eventualities, places, times and other possibly interesting parts of the discourse would not change anything in the general principles of representation – they would all be referents specified by the relevant properties (such as duration, distance, concatenation) and associations among them.

The condensed introduction of MADR above should serve its intended purpose in this paper: to present how a (segment from a) flat structure like this can be turned into a hierarchical one typically proposed as the (morpho-)syntactic structure of a natural language sentence. This latter step is discussed in the next section. Observe that MADR structures are non-oriented and non-hierarchical. None of the referents is in any sense structurally subordinated to any other, and the interpretation of none of the referents has to precede the interpretation of any other. As it will turn out, the structure of MADR is multi-hierarchical, or at the level of any single association – it is ambi-hierarchical.

Some aspects of MADR

One of the major properties of MADR is that it is not hierarchical in the strict traditional sense. The hierarchical relation between two referents is not uniquely determined. Since the web is in some sense flat – all its nodes belonging to the same level – a large number of different hierarchies can be derived from its organization. None of these hierarchies is more natural, or in any way more prominent than the others. In fact, in order to have any hierarchy, we first have to determine one referent from which to start as the hierarchically highest. Once this is done, all the referents directly associated to this chosen one through their properties will belong to the second level of the hierarchy, each forming its own sub-hierarchy; all their associates will belong to the third level, and so on. This resembles a mobile which lies flat on some surface and can be taken by any of its threads. Whichever thread taken, the others will belong to the second, hanging level. Choosing any of the referents of a MADR results in a different hierarchical configuration. Hence, the mapping between hierarchically structured natural language expressions and MADR is many to one: many expressions differing only in the hierarchical organization map to one and the same segment of MADR. This enables a direct update of all the referents that a linguistic expression informs about irrespective of its hierarchical organization and hence also independent of which referent is its subject.

Discourse referents are defined via discourse roles. A referent corresponds to a set of properties associated with one discourse role. The set of properties provides the description of the referent. These properties are associated to other properties, describing other referents, which makes the description of each referent highly dependent on the entire discourse, and especially on the descriptions of referents that are directly associated to it.

Units that can be recognized in the discourse are the following: properties and associations as atomic elements, referents as a derived type, and different segments of the MADR web, containing a number of mutually associated referents. The denotation of one complete nominal, verbal, or clausal expression corresponds to one referent, but usually, in describing this referent, it includes a number of other referents associated to the targeted one, in line with the mobile metaphor above. Natural language expressions can be observed as instructions for the retrieval of the discourse representation. Their hierarchy determines the sequence of moves in this retrieval: it starts with the referent the expression corresponds to (the ball in the example in (4)), and specifies this referent through its discourse role (encoded by markers like demonstratives and by the position and prosody of the expression in a broader linguistic context) and its description. The description may include other referring expressions, usually those that have recently been integrated into the discourse, or for some other reason easier to be mapped (the left corner and the desk in (4)). The interpretation of a linguistic expression equals its full mapping onto a segment of MADR and the identification of the referent targeted by the retrieval procedure (the hierarchically highest one in the linguistic expression: the ball in (4)). In communication, this process, when targeting the sentential level, usually results in a mismatch between the located segment of MADR and a special, intonationally and structurally marked part of the denotation of the linguistic expression. In such a case, this mismatched part is incorporated into the discourse representation as new

information, enriching it by a relatively small segment (usually one new referent, an object or an event, with the properties associating it to other referents).

(4) [that red little ball [in the left corner], [behind the desk]]

MADR is clearly non-linear, and in fact without a strict linearization strategy. Linearity is not part of its nature in any interesting way. Linear structures in language are purely grammatical, and emerge at the other side of the interface. Linear structures emerge either after the application of the linear computation on MADR (the retrieval procedures, or some other, syntactic, linear computations), or for some even later, at the interface with phonology. This property of MADR, non-linearity, and its web-like organization, needs some motivation. To argue that linearity is well explained as an import of the computational module, or of phonetic aspects of language, does not exclude that other modules might also involve some linear structures. At the same time, all that we know about language and the related phenomena like discourse, came by our observation of linear structures of linguistic expressions, from phrases, to sentences, to texts. Our only window into the structure of the discourse is linear. Possible arguments for nonlinear structures at some deeper level could come from a) observations of syntactic processes that look like strategies of linearization applied to non-linear structures and b) independent, probably purely conceptual, arguments why non-linear structures would better serve the purpose of levels of representation such as the discourse. I present two speculations, hinting towards these two directions.

One of the typical properties of the human language syntax relates to the displacement phenomena observed in virtually all languages of the world. In many languages, the direct object, realizing the affected participant in an event, canonically appears in a position after the verb, and hierarchically in a very low position. When a wh-question is asked targeting this participant, the wh-word, taking the same role, appears in the beginning of the sentence, ad hierarchically in a very high position. Both theoretical and psycho-linguists argue that the same wh-word in such cases receives interpretation in two positions, the very high one for its wh-component, and the very low one for its participant role. Yet, in the sentence, it surfaces only once, in only one of these positions.

(5) a. John ate an apple.
b. What did John eat?
wh did John eat THING
|_____|

The need to have the same element in two positions and then introduce complex rules of deleting one of its instances under subtle locality conditions indicates that syntax is operating on a different kind of structure in order to make it phonologically realizable. In a web-like representation, the double interpretation that the same element receives in (5) would be easy to achieve with only one instance of this element, simply by attributing to it two different properties, associated to two different other referents. Displacement phenomena, otherwise difficult to motivate in grammar, receive explanation in terms of a strategy to linearly represent two distinct properties (with their associations) of one and the same referent.

A conceptual argument in favour of the web-like representation of MADR, compared to the list-like representations like the file-card metaphor, is highly economical. In the file card metaphor, taking a list of propositions about the same

referent, each of the propositions needs to involve a representation of this referent, and probably also of a number of other referents. What is worse, the information carried by some propositions would be represented several times, requiring one proposition for each referent that it concerns. This abundance of material is avoided in MADR. Each referent is represented only once, and each different bit of information corresponds to only one segment in the representation. At any point, a relatively simple retrieving engine could produce the full list of referents and propositions from a significantly more economical MADR.

Finally, lists have not been part of any serious proposal of how the brain works, at least of none that I know of. Webs, on the other hand, present one of the typical structures used to model neuronal activities. Moreover, biological and computational theories of spatial cognition propose a very similar model in which cognitive maps are flat webs, but their computation is hierarchical, and can start from any of the nodes in the web (Voicu 2003). I am not claiming that MADR represents some kind of neuronal reality. Yet, if the neuronal reality behind the discourse is well represented in terms of a web-like structure, the translation from that structure to MADR and the other way around should be easier than with a list-like representation.

Embedded discourse representations

The descriptive formula in (2) (Discourse = {R | R={Prop | Assoc (Prop, Prop') & Prop' ∈ R'}}) defines a discourse as a set of (interconnected) referents. This implies that the discourse is to be seen as a second derived type, in addition to referents. Two interesting immediate questions are: Is it possible that one interlocutor operates over two or more different discourse representations in the same communication situation? and Can discourse representation form more complex structures, and what these structures are? I argue that the answer to both questions is positive, and that the structure involving more than one discourse presents discourse as well, i.e. that discourse representations embed into other discourse representations.

Many empirical facts of language, communication and cognition require a simultaneous representation of more than one discourse. If we take that the discourse includes all the presupposed material and all the material that has been communicated within a certain communication situation, then a number of situations involve more than one such object. For instance, if a group of five people takes part in communication, there usually is a body of information available to a subgroup of two or three interlocutors. Moreover, communication often makes use of such facts: sometimes a sentence would only be incorporable into the discourse of one of the subgroups. If a speaker produces two adjacent utterances, such that one updates the discourse of the whole group and one the discourse of one of the subgroups, the speaker must be dealing with two different discourse representations. It is possible that these different discourse representations are independent of each other. However, the discourse of the maximal group is always identical to a segment of each of the subgroups' discourses. It is at least a more economical representation would be the one in which the different discourse representation 'intersect' and even perhaps share some of the roles.

The number of different discourses relevant for a certain communication situation is not unpredictable or random. Every subset of the aggregate group of interlocutors,

including singletons, has an own discourse. Evidence for this is abundant. For instance, specific reference is intuitively defined as the property of referring to a particular entity which can be identified, and which cannot be substituted by any other. Observe the sentence in (6) in a communication situation which involves both the speaker and John as interlocutors.

(6) John bought a certain TV.

In this case, there has to be a discourse in which the referent of the expression *a certain TV* has not been represented so far. It is either only John's discourse, or also the one shared by John and the speaker, in which the particular referent is old. The possibility that only John knows exactly which TV he bought implies that there is a discourse representation for each of the interlocutors, in addition to all the shared ones.

But consider now the situation in which John is not among the interlocutors. It is still possible that only John can identify the exact TV that he bought. This means that not only all the interlocutors, but in fact probably all the animate referents in the discourse involve their own discourse representations, and hence have to be represented together with an own, embedded, discourse representation.

This view gains support from the phenomenon traditionally called the *de se* vs. *de re* readings distinction (e.g. Anand 2006). The sentence in (7), out of the blue, sounds somewhat odd: how could a man have beliefs, rather than knowledge, about his own profession? However, in a context in which the old man is an amnesiac, found on the street with no documents, the sentence suddenly makes sense.

(7) The old man believed that he was a painter.

But in this context, the old man thinks of himself as of a different person. This is even clearer in (8), for a context in which John listens to recorded voices, not knowing that his is among them, and appears to like the least the recording of his own voice.

(8) John liked his own voice the least.

The important fact here is that there is a discourse in which the voice is not specified as John's, and another discourse in which this information is available. The former discourse is most probably John's individual discourse, and the latter the one of the speaker of the sentence in (8). We can even replace *John* with *a cat*, with the discussion above still holding, which means that the relevant property is indeed animacy. This opens the way towards explaining the prominent nature of animacy in grammar: animate referents in every discourse are special because each of them introduces an own embedded discourse into the relevant macro-discourse.

This view allows us to formulate the formal definition of communication in (9).

(9) Communication presents a series of subsequent transfers of segments of the individual interlocutors' discourse representation into the discourse representation shared by the entire relevant group of interlocutors.

Another level requiring the embedding of discourse representations is that of modality. The domain of every modal operator can be viewed as a certain whole. This whole has all the properties of a discourse representation and at the same time is part of the bigger discourse in which the modal operator appears. It is in fact possible to put an entire communication situation under a modal operator (*It could be that,*

after the dinner, John said that they should leave, but Mary thought it was too early and told him to wait.), in which case the discourse nature of the modal domain is uncontroversial.

The need to embed discourse representations is not straightforwardly handled by MADR, but there are ways how it could be done. One possible solution is to introduce a special type of properties which are direct members of discourses (hence embedding them under the referents whose properties they associate with). I leave this aspect of MADR for future work.

Conclusion

Linearity in the discourse, as well as in language, might be a consequence of the syntactic and phonological aspects. On the semantic side, the discourse may well be represented in non-linear terms. Such representation has the important advantage of allowing for a single utterance to simultaneously update all the referents whose representations are affected by the contributed information. It is also more economical than list-based representations and structurally closer to the patterns observed in the neurological domains. An open-bottomed decompositional approach that indeed might result in very complex representations, pays off by a very high level of precision and flexibility. The paper sketched a segment of a possible technical implementation of the discussed ideas, which needs to be developed and formalized further before its potentials and problems could be evaluated.

References

Anand, P. 2006. *De de se*. Doctoral Dissertation, MIT.

Bierwisch, M. 1983. Semantische und konzeptuelle Repräsentation lexikalischer Einheiten, In R. Ruzicka and W. Motsch (eds.), *Untersuchungen zur Semantik*, 61-99. Berlin: Akademie-Verlag.

Carstairs-McCarthy, A. 1999. *The Origins of Complex Language: An Inquiry into the Evolutionary Beginnings of Sentences, Syllables, and Truth.* Oxford University Press.

Dowty, D. 1991. "Thematic proto-roles and argument selection." *Language* 67.3: 547-619.

Heim, I. 1982. *The Semantics of Definite and Indefinite Noun Phrases*, Doctoral dissertation, UMass-Amherst.

Herweg, M., and C. Maienborn. 1992. Konzept, Kontext, Bedeutung – Zur Rolle der Zwei-Ebenen-Semantik in einem Modell der Sprachproduktion, In M. Herweg (ed.), *Hamburger Arbeitspapiere zur Sprachproduktion* 9, 7-31.

Van Valin, R. D., Jr. & R. J. LaPolla. 1997. *Syntax: structure, meaning & function.* Cambridge: Cambridge University Press.

Voicu, H. 2003. Building and using a hierarchical representation of space, *Neural Networks* 20-24 – *Proceedings of the International Joint Conference on Neural Networks* vol. 2: 1334-1339.

Boban M. Arsenijević

МУЛТИХИЈЕРАРХИЈСКА РЕПРЕЗЕНТАЦИЈА ДИСКУРСА

Сажетак

Заступан је став да у концептуалном домену, као и у неуро-когнитивном – доменима у којима сагледавамо операције организовања и ажурирања репрезентације дискурса, нелинеарне структуре носе виши генеративни и меморијски капацитет од линеарних, које су карактеристичне за фонетски и друге сензори-моторне домене. Уместо традиционалних линеарних модела репрезентације дискурса, рад предлаже мултихијерархијски семантички формализам, базиран на само два семантичка типа: својствима и референтима (као скуповима својстава), где су својства дефинисана као релације између референата и других својстава. Расветљене су неке од предности, али и неки од недостатака оваквог заснивања модела.