Verbs and Events in Discourse: A Synergy

Alexandros Tantos

Aristotle University of Thessaloniki
atantos@gmail.com

Abstract

Verbal semantics has been shown a useful resource for reconstructing not only the semantics of whole sentences, but also of inferring relations between larger pieces of text. As Asher and Lascarides (1995), Danlos (2001a,b) and Tantos (2008, 2009) point out, lexical items have the power to simplify discourse inference and explain the ease with which we resolve ambiguities and construct the meaning of discourse. Based on the well-studied case of causative verbs in syntax-semantics interface, this paper aims to prove that it is not only a desired strategy to include lexical semantic knowledge in modeling discourse inference and interpretation, but also essential to include such lexically triggered information within the analyses of discourses.

1. Introduction

Since the first dynamic semantic theories (File Change Semantics by Heim 1981 and Discourse Representation Theory (DRT) by Kamp and Reyle 1993), it was made clear that content and structure in discourse go hand in hand within a hierarchical model of semantic representation and interpretation. The establishment of structural constraints in the update of discourse information paved the way for providing descriptive adequateness for a number of semantic phenomena, such as plurals, negation and pronoun resolution. One of the most successful tools and diagnostics for verifying or falsifying the reality of hierarchical models of discourse structure is anaphora resolution.

(1)   a. John's roommate$_i$ came to the party yesterday. He$_i$ brought his own beer.
        b. * If his roommate went to a party$_i$ yesterday he must have been drunk. It$_i$ was crazy.
        c. John$_i$ left yesterday and if he$_i$ went to the opera he must have been very satisfied.

In (1a) the two-sentence discourse allows *John's roommate* to be the antecedent of the pronoun in the second sentence.[1] (1a) is the easy case of resolving anaphoric elements, since the semantics of the discourse can be seen as a simple conjunction of the contents of the two sentences rendering all its content available for building a coreference relationship. Therefore, if the pronoun *it* denotes the variable *x*, then the assignment function *g* has no other option (with respect to its gender assignment) than to assign the value of the constant *s* brought by *his roommate* to the variable x, i.e. *g(x)=s.*

However, in (1b) things are different. The pronoun *it* and the indefinite *a party* do not corefer with no apparent reason prima facie. The reason why this semantic misfire happens lies in the semantic constraint brought by the fact that the potential antecedent *a party* occurs within the conditional of the first sentence. Entities referred or *discourse referents* within conditionals have been shown not to be available for a coreference relationship with discourse referents outside the conditional. on the contrary, *discourse referents* within conditionals have *access* to referents outside of it as in (1c). This asymmetry as well as a number of other grammatical constraints relevant to anaphora resolution, such as quantificational operators, plurals and temporal phenomena, show that grammar affects and is affected by the hierarchical structure of discourse. In fact, the violation of the constraints leads to discourse incoherence in a strict interpretation parallel to that of ungrammaticality. Semantic structure is then the first and most obvious factor for imposing a hierarchical structure to discourse for expressing connections between bits of information within discourse.

Soon after entering formal semantics' machinery for interpreting discourse, influential theories of discourse semantics, such as DRT, acknowledge the presence of pragmatic factors that structure and shape partially but substantially the notion of discourse coherence. For instance, the function of using the second sentence in (2) is to *explain* the event described by the verb in the first sentence. Following textual order, the *pushing* enters discourse after the *falling*. The occurrence of *falling* sets a temporal anchor or reference point after which the next-mentioned event initiates. However, the default interpretation is that the *pushing* occurred before the *falling* and, thus, the semantic content of the discourse is not simply a conjunction of the contents of its parts in a compositional way similar to that of the semantic composition of the majority of single sentences.[2]

(2)   John fell. Max pushed him.

---

1 I assume here that coindexing expresses the coreference relationship for the sake of the examples in this paper without implying any one-to-one connection between the syntactic mechanism of binding and the semantic coreference.
2 Asher and Lascarides (2003) offer a rich and well-established inventory of substantial reasons why dynamic semantics is not enough for interpreting discourse.

Therefore, the rhetorical structure of the discourse is an additional parameter that provides discourse its hierarchical flavor. In (2), *explanation* is the rhetorical function that imposes the reverse temporal order of the events in the discourse.

Among the existing computational and pragmatic theories of discourse structure, only Rhetorical Structure Theory (RST, Mann and Thomason 1987) and Segmented Discourse Discourse Representation Theory (SDRT, Asher 1993, Asher and Lascarides 2003) take on board the idea that the representation of discourse information should include a well-defined inventory of rhetorical relations. Particularly, RST and SDRT implement discourse coherence as directed graphs with edges connecting Elementary Discourse Units (EDUs) or else discourse segments with rhetorical relations as representational devices. Moreover, only SDRT offers a formally precise framework that allows intrasentential information to interact with intersentential ones. Specifically, SDRT's discourse structure, SDRS, is a triple (A, F, Last), where:

- A is a set of labels that represent the EDUS, which represent the primitives of discourse or else the discourse terms of our formal apparatus and function as arguments to discourse relations.
- Last is a label in A (the EDU label that represents the content of the last clause inserted to the discourse logical form).
- F is a function that assigns each member of A a member of a formula of the SDRS language, which includes formulas of some version of dynamic semantics (DRT, DPL, Update Semantics, MLTT, etc.).

Due to the formal semantic machinery that SDRT provides, which is quite illuminating for our purposes of associating lexical with discourse semantics, in the rest of the paper, I will first provide a crash course to the main factor that regulates discourse hierarchy, the principle called Right Frontier Constraint (RFC) as implemented within SDRT, in order to assess the validity of the examples and the conclusions of the paper. Then, I will show why it is not only more effective, but also essential to allow rhetorical relations to interact between EDUs as well as lexically triggered discourse units. The last part of the paper is devoted to a new view of how segmentation of discourse information works and provides an insight as to what criteria are necessary to define EDUs.

2. RFC and Discourse Relations

The Right Frontier Constraint is originally proposed by Polanyi (1988) and is one of the most well-documented constraints in literature of formal discourse semantic and pragmatic theories. RFC is deemed responsible for the way discourse update is built and discourse anaphora is resolved. In order to understand the importance of RFC for explaining the hierarchical flavor of discourse, it is essential to grasp the distinction between two types of discourse

relation, subordinating and coordinating ones. Briefly, the idea behind this distinction is that by uttering the current segment the author-speaker either provides information in the same direction with respect to a prior segment within a range of different utterance modes, such as explaining, elaborating or commenting, (subordinating relations) or entirely shifting the topic of discourse to new directions with various ways, such as narrating, providing a background or a result of the prior segment (coordinating relations).

The type of rhetorical relation is the most important part of RFC. Schematically, the difference between the two types with respect to RFC is reflected within graphical representations. The nodes of a SDRS graph represent its labeled constituents and its edges represent the discourse relations between these constituents. Thereby, downward edges express subordinating relations and horizontal edges express coordinating ones. The downward direction in the case of subordinating relations signals the degree of granularity of the information provided. An important restriction is that two nodes in a graphical representation cannot be connected using both a subordinating and a coordinating relation.

For instance, in (3), a simple three-sentence discourse, contains five EDUs and the SDRS graph is depicted in figure 1. π2 and π3 *elaborate* on the visit to the ophthalmologist mentioned in π1.[3] On the other hand, π4 and π5 do not provide any more detail about the event described in π1. These EDUs denote a shift on the topic and they continue the discourse in a different direction, in the sense that they describe new events that happened one after the other, the *driving* and the *drinking*. π4 builds a *narration* relation with π1 and not with π3 or π2 and, hence, since *narration* is a coordinating relation, it is connected with a horizontal edge with π1.[4] Accordingly, π5, being related *narratively* with π4, is connected with a horizontal edge with it.

(3)   Yesterday evening I visited an ophthalmologist (π1). He diagnosed a small eye problem (π2) and wrote me a prescription (π3). Then, I drove back home (π4) and drank a beer (π5).
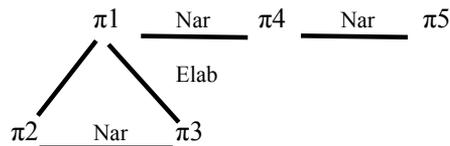
The SDRS graph of (3).

---

3 EDUs are usually represented by the Greek letter π. Throughout the paper this convention is used.

4 According to Asher and Vieau (2005), narration and elaboration are the prototypical coordinating and subordinating discourse relations respectively.

Asher (1993) implements RFC by inserting a rule defining the right frontier. The EDUs that lie on the right frontier of the discourse graph built so far are available for attachment; in other words the update of information in the discourse allows new information to either attach to the last entered EDU β in a discourse structure or to some other constituent γ such that (β,γ) is in the transitive closure of the subordination relation. More formally, the right frontier is defined in SDRT as the set of available nodes for attachment falling under the following possibilities:

1. The label α = LAST;
2. Any label $\gamma \geq D^* \alpha$, where $\geq D^*$ is defined recursively:
   a) $R(\gamma,\alpha)$ is a conjunct in F(l) for some label l, where R is a subordinating discourse relation;
   b) $R(\gamma,\delta)$ is a conjunct in F(l) for some label l, where R is a subordinating discourse relation and
   F(δ) contains as conjunct R'(δ',α) or R'(α,δ'), for some R' and δ'; or
   c) $R(\gamma,\delta)$ is a conjunct in F(l) for some label l, where R is a subordinating discourse relation and
   $\delta \geq D^* \alpha$.

Attachment of new information is directly influenced by the number and placement of available EDUs in the right frontier. At the same time, availability of attachment points on the right frontier should allow no more and no less than the EDUs that correspond to the context, since otherwise its definition would lead to over- and/or undergeneration in terms of discourse structures.

3. Discourse Segmentation and Anaphora Resolution

The majority of formal discourse semantic theories, such as SDRT, acknowledge the essence of lexical semantic knowledge for inferring and interpreting discourse. However, the need for including lexical semantics should be established within and by a solid theory based on real language examples on a case-by-case study. For our purposes, anaphora resolution is perhaps the most reliable test and criterion by which one can enter the discussion of how natural language speakers segment utterances for encoding semantic relations between bits of information in the discourse. In other words, following Asher (1993), I intend to cash out the distinction between real and natural language metaphysics by assuming that pronouns refer to real entities of natural language metaphysics with no commitment as to the model interpretation of those entities. Keeping that in mind one could factor out possible antecedents to anaphors as real entities independently of their grammatical form.

The immediate repercussion of this idea is that it is possible to focus on semantic entities that would do not necessarily correspond to syntactic forms. The appeal to semantic entities that are part of our daily discussions is not only desired but also necessary in order to provide simple analyses close to linguistic reality, since otherwise the coherence of smaller or larger discourses cannot be explained. Let us consider an illuminating example below:

(4)   Yesterday John broke the carafe (π1). Its pieces were lying all over the floor (π2) and it was such an elegant carafe (π3). He was deep in his thoughts (π4) while he was walking up and down the room (π5) and hit it with his knee (π6).

In (4) it seems we are able to pick up and talk about parts of the event of breaking.[5] Moreover, these parts are well-defined and they are related to the cause and result of the breaking event, they are not arbitrary spatiotemporal bits of information and they influence discourse coherence. The EDU π1 refers to an event of breaking with an agent and a theme. π2 *elaborates* on the state of the broken carafe and π3 *comments* on a property of the theme in π1, before the breaking took place. Furthermore, both verbs of π2 and π3 are in the same past tense and there is no obvious temporal marking that differentiates between the *before* and *after* the attainment of the broken state. π4 and π5 provide *background* information and π6 *elaborates* on the cause of the breaking state. Furthermore, π6 is related to π1, elaborates on the cause and not on the result of the breaking. If one does not take for granted the presence of EDUs triggered by lexical items, the verb '*break'* in (4), there is no obvious way of explaining:
   a) why we take a reading in which π2 refers to a state initiated well after the initiation of the state of John b*eing-deep-in-thoughts* in π4 and the state of *walking-up-and-down-the-room in* π5,
   b) why we interpret π2 *elaborating* on π1 and at the same time π4 and π5 provide *background* information on π1 while π2, π4 and π5 being states related to π1, and
   c) why we can resolve the anaphor '*it'* in π6 with *the carafe* in π1, when RFC is violated.[6]

An immediate answer to these issues and the ease with which we infer and interpret events in various discourses would be to include some Neo-

---

5 We are also able to talk about properties of the participants of an event. However, this has always been considered part of discourse processing of salient entities.
6 It violates RFC, since EDU 6 cannot be related to 1 due to the fact that it is not in an accessible attachment site. For more on RFC, cf. Asher and Lascarides (2003).

Davidsonian event-based theory following Parsons (1990) that separate the causing part of an event, as a subevent that precedes the event of a result state. However, this approach does not take us much further for the following reasons: a) one should appeal again to subevents as part of the analysis, and since subevents are ultimately events in their turn , this leads us to problems related to their identity conditions, as Tantos (2011) notes; b) I see no apparent way of formally encoding relations between events that would cover rhetorical relations. Questions relating to issues of elaborating on an event, providing background to it, or explaining it on the basis of another event, are totally unclear in terms of model interpretation. Discourse relations connect bits of related information in a specific context of usage.

The solution I suggest is to include lexical labels in the repository of discourse theories and assign them a special status on account of their implicit presence. By implicit presence I mean that their presence is not marked by any sort of explicit marker and is explicated only when discourse inference and interpretation is pursued. Following work by Danlos (2001a, b) and mainly by Tantos (2008, 2009), I claim that lexical EDUs are present in discourse inference and carry typed information, but these EDUs should not be constrained in the same way that other EDUs are. In (4) the verb '*break*' inserts two lexical EDUs, 1a and 1b, which are not subject to the RFC and, thus, they are accessible to EDUs $\pi 2$, $\pi 4$ $\pi 5$ and $\pi 6$. Moreover, $\pi 1$ is still present and accessible to other EDUs following the standard rules of discourse inference. That is why $\pi 3$ has access to it and comments on *the carafe* independently of the cause and result of breaking. If $\pi 1$ were not present, and all there was were the lexical EDUs, it would be incoherent to talk about *the carafe* with no mention to the event of breaking (cause and result). In other words, verbs insert lexical EDUs without disturbing the segmentation of the discourse.

Discourse segmentation is a difficult task for any formal semantic theory and any system of computational semantics. The main reason lies on the fact that until now the majority of attempts supports a one-to-one correspondence between surface forms such as sentences and propositions. Only recently, Schlangen (2003) and Schlangen, D. and A. Lascarides (2003a, b) consider a formal approach in which surface utterance forms may correspond to complex EDUs. Discourse segmentation, then, has the main burden for a whole class of inferences when causative verbs are mentioned in the discourse. Lexical EDUs function as default templates for inferring connections with other utterances in the discourse and have a conceptual status clearly intermediating between linguistic and extralinguistic knowledge. It is this level of description that one needs to investigate more in order to decode inference patterns on the one hand and the limits of linguistic theorizing in discourse.

4. Conclusion

By acknowledging the existence of lexically triggered information, we are in the position to explain in a natural way why discourses are interpreted easily by human interpreters without resorting to heavy inference machinery about possible event connections. In this paper I showed that a) the current approach to discourse segmentation along with its definition of attachment availability is not able to describe the complexity of discourse connections within a single example that includes a causative verb and that b) one should not be based on event-based theories for inferring and interpreting discourse connections, since these are semantic entities for sentential interpretation and do not directly affect discourse ineference and interpretation. Moreover, one needs to assume lexical EDUs that have clearly a conceptual status and serve default inferences in discourse.

References

Asher, N. 1993. *Reference to Abstract Objects in Discourse*. Dordrecht: Kluwer.

Asher, N. and A. Lascarides. 1995. Lexical disambiguation in a discourse context. Journal of Semantics 12(1):69–108.

Asher, N. and A. Lascarides. 2003. Logics of Conversation. Cambridge: Cambridge University Press.

Asher, N. and L. Vieu. 2005. Subordinating and coordinating discourse relations. *Lingua*, 115(4), 591–610.

Danlos, L. 2001a. Event coreference between two sentences. In H. Bunt and R. Muskens, eds., Computing Meaning 2. Kluwer Academic Publishers.

Danlos, L. 2001b. Event coreference in causal discourses. In P. Bouillon and F. Busa, eds., The Language of Word Meaning. Cambridge University Press.

Heim, I. 1982. The Semantics of Definite and Indefinite Noun Phrases, University of Massachusetts: Ph.D. dissertation; published 1989, New York: Garland.

Kamp, H. and U. Reyle. 1993. From Discourse to Logic. Dordrecht: Kluwer Academic Publishers.

Parsons, T. 1990. *Events in the Semantics of English*. MIT Press.

Polanyi, L. 1988 A formal model of the structure of discourse. *Journal of Pragmatics*, 12, 601–639.

Schlangen, D. 2003. A Coherence-Based Approach to the Interpretation of Non-Sentential Utterances in Dialogue. PhD Thesis, School of Informatics. University of Edinburgh.

Schlangen, D. and A. Lascarides. 2003a. The Interpretation of Non-Sentential Utterances in Dialogue, in Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue, Sapporo, Japan, July 2003.

Schlangen, D. and A. Lascarides. 2003b. A Compositional and Constraint-Based Approach to Non-Sentential Utterances, in S. Müller (ed.) *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar*, Michigan, USA, July 2003.

Tantos, A. 2008. *Computing Events in Discourse: A Case Study Involving Light "have"*. PhD Dissertation, *Konstanzer Online-Publikations- System (KOPS),* Konstanz.

Tantos, A. 2009. Lexically Triggered Discourse Relations. In *Selected Papers from the 18th International Symposium on Theoretical and Applied Linguistics*, A. Tsangalidis (eds.), Dept of English Language and Literature, Aristotle University of Thessaloniki.

Tantos, A. 2011. Κειμενογλωσσολογικοί Παράγοντες στη Διδασκαλία του Κλιτικού Αναδιπλασιασμού της Ελληνικής σε Φυσικούς Ομιλητές Γερμανικών Γλωσσών. Στο *Μελέτες για την Ελληνική Γλώσσα* 31, 485–495, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης: Θεσσαλονίκη.