How to infer temporal relations in discourse?

MOESCHLER, Jacques

Abstract

This paper is devoted to the following question: how can readers and listeners infer temporal relations in discourse? One of the basic inferential task in utterance and discourse interpretation consists in inferring time direction, that is, in determining the time of the event relative to the thread of discourse. In this paper, I reduce time direction to two temporal relations: forward inference (FI) and backward inference (BI). I show that the computation of directional inference is neither the result of principles of discourse, nor the consequence of discourse type, but the interaction of information coming from different sources, that is, contextual information and linguistic information.

Reference

MOESCHLER, Jacques. How to infer temporal relations in discourse? In: Aurnague, M. ; Bras, M. ; Le Draoulec, A. & Vieu, L. *First International Symposium on the Exploration and Modelling of Meaning - SEM-05*. Toulouse : IRIT, 2005. p. 133-142

How to infer temporal relations in discourse?

JACQUES MOESCHLER

Department of Linguistics, University of Geneva, Jacques.Moeschler@lettres.unige.ch

ABSTRACT

This paper is devoted to the following question: how can readers and listeners infer temporal relations in discourse? One of the basic inferential task in utterance and discourse interpretation consists in inferring time direction, that is, in determining the time of the event relative to the thread of discourse. In this paper, I reduce time direction to two temporal relations: forward inference (FI) and backward inference (BI). I show that the computation of directional inference is neither the result of principles of discourse, nor the consequence of discourse type, but the interaction of information coming from different sources, that is, contextual information and linguistic information.

1. INTRODUCTION

In recent works on discourse semantics (DS), one major step has been passed while distinguishing two types of knowledge active in discourse interpretation: linguistic knowledge (LK) and world knowledge (WK). In DS, there is a general agreement about the following fact: while LK is defeasable, WK is not. This principle is a general principle of pragmatic analysis, stating that LK is weaker than WK. If we try to understand this principle relatively to conversational implicatures, the parallelism is striking: implicatures are inferences that can defeat literal and compositional meaning, that is, LK. In discourse semantics (as for instance in Asher 1993, Lascarides & Asher 1993, Asher & Lascarides 2003), discourse relations (DR) are the by-product of WK and LK. For instance, in narrative discourses, LK allows to infer Narration, whereas WK can defeat it, as shown in (1) and (2):

(1) Max pushed John. He fell.
(2) John fell. Max pushed him.

The standard interpretation of (1) is Narration (forward inference, FI), whereas the forward reading in (2) is generally defeated in favour of the Explanation or backward inference reading (BI).[1] In other words, if no information coming from WK cancels LK, we are allowed to conclude that the events described in discourse have to be ordered as the discourse processes. Narration is thus the case when no conflict occurs between LK and WK, whereas Explanation is the case when the Narration default reading crashes.

One of the major contribution of Post-Gricean radical pragmatics, as Relevance Theory (RT), is the distinction between conceptual information (CI) and procedural information (PI) (Blakemore 1987, Wilson & Sperber 1993a), and the parallelism between CI and PI and semantic and pragmatic or truth-functional and non-truth-functional information. While CI encodes information about the state of affairs, concept-type expressed by the lexical item linked to the concept, PI is information about how to treat CI. In narrative discourse, it is crucial both for access to representation of eventualities described by utterances and for information indicating how to process between these eventualities. For instance, the

---

[1] English allows both backward and forward readings for (1) and (2), whereas French Passé Simple forbids it:

(i) Max poussa Jean. Il tomba FI^{ok}, BI*
(ii) Jean tomba. Max le poussa. FI*, BI^{ok}

Cf. Moeschler (2003) for a general description of tense combination and directional interpretations.

computation of directional inference (DI) typically results of applying instructions encoded within PI on representation of events.

This picture is nevertheless too simple. On the one hand, DIs are not simply the by-product of LK and WK, in the sense that LK can defeat WK as much as WK can defeat LK. On the other hand, PI is not the only type of information triggering DI: so can CI, as the parallel French examples in Passé Composé shows it:

(3)	Max a poussé Jean. Il est tombé	*Max pushed John. He fell*	FI^{ok}, BI*
(4)	Jean est tombé. Max l'a poussé	*John fell. Max pushed him*	FI*, BI^{ok}

The question is now the following: how does DI occur in discourse if it is not the simple result of WK on LK of PI or CI? What we know about linguistic data is that DIs are not inferred randomly and that some types of discourse are more optimal than others relative to DIs. In other words, we need a complete explanation that can interface linguistic with non-linguistic information and utterance with discourse interpretation.

2. PRAGMATIC OR DISCOURSE EXPLANATION?

The first point to make explicit is the following: do we need any type of *discourse knowledge* (DK) to process DIs in discourse? This question must be seriously asked, even if it has received a consensual positive answer (see Asher 1993, Caenepeel & Moens 1994, Dowty 1996). The consensual positive answer (cf. Vet 1995 for a cognitive motivation) consists in claiming that DK is crucial to the determination of DI. Thus Narration (FI) is highly preferred in narrative discourse, whereas Explanation (BI) is the case in non-narrative discourse (Labov & Waletztky 1967). But this assumption has heavy consequences: it supposes that while processing a narrative discourse, we have a DK such that we know exactly what we have to know, that is, in order to move time forward, we know that we do so because we are processing a narrative discourse. DK is thus a crucial pre-requisit, and if it can be proved that no such knowledge as DK is required, then the explanation cannot presuppose any discourse rule or principle.[2]

What could be an alternative to a discourse-oriented explanation to the computation of DI? Linguistic explanation would suppose that the only procedural information encoded in tenses would be responsible for the triggering the good DI. This solution, proposed for instance in Kamp & Rohrer (1983), meets too many counterexamples to be considered satisfying.[3] The only solution is thus pragmatic, that is, a solution combining linguistic information (and more precisely CI and PI) and contextual information. No specific discourse rule or principle is needed, as we shall see. What we need on the contrary is a theory on how linguistic information combines within contextual information. We claim that Relevance Theory provides the necessary theoretical background.[4]

---

[2] This is the thesis advocated in Reboul & Moeschler (1998), in which discourse comprehension is explained as a non-demonstrative inference process including hypothesis on communicative and informative intention attributed to both utterance and discourse (respectively local and global communicative and informative intentions).

[3] In Kamp & Rohrer (1983), French Passé Simple encodes Forward Inference, French Plus-Que-Parfait encodes Backward Inference, and French Imparfait encodes Inclusive Inference.

[4] In Moeschler (2000), we claim that DI are not Gricean implicatures (neither conventional nor conversational), but contributions to the explicatures of the utterances (they contribute to the truth value of the utterance).

3. THE MODEL OF DIRECTIONAL INFERENCES (MDI)

In MDI (Moeschler 2000), we assume that two main hierarchies organise the relevant types of information active in processing DI: contextual information vs. linguistic information on the one hand, and procedural vs. conceptual information on the second. Among procedural information, we further distinguish scopes: propositional vs. morphological. We obtain thus the following hierarchy of information:

A. Contextual information is stronger than linguistic information.
B. Procedural information is stronger than conceptual information.
C. Propositional procedural information is stronger than morphologically incorporated procedural information.

These three principles organises the linguistic system and the principles of its use. Principle A means that a contextual assumption can defeat a DI inferred on the basis of the sole linguistic information. Principle B states that tenses are stronger than concepts, for instance causally related concepts. Finally, principle C claims that all functional materials in C or higher than T in the syntactic tree are stronger than tenses located in the head of TP.[5] Let us illustrate these principles.

3.1 Principle A

In (5) and (6), time direction comes from the projection of a causal rule (7) onto a contextual assumption (8), the normal contextual implication being (9):

(5)	Max a poussé Jean. Il est tombé	*Max pushed John. He fell*	FI^{ok}, BI*
(6)	Jean est tombé. Max l'a poussé	*John fell. Max pushed him*	FI*, BI^{ok}
(7)	(PUSH, x, y) CAUSE (FALL, y)		
(8)	If Max pushed John, then John fell.		
(9)	John fell after Max pushed him.		

Now we can imagine another contextual assumption, for instance (10). (5) and (6) would then have respectively a BI and a FI reading, allowing the new contextual implication (11):

(10) Max pushed John and then John fell.
(11) Max pushed John because John fell.

These readings are certainly less probable than the standard forward (5) and backward (6) ones, but could occur in appropriate contexts. Thus, Principle A does not explain only standard cases, but also non-standard ones.

3.2 Principle B

French Passé Simple (PS) and Plus-Que-Parfait (PQP) are the best illustration of Principle B. In examples (12) to (15), procedural information given by tenses either confirms or defeats conceptual information given by the push-fall causal rule:

(12)	Max poussa Jean. Il tomba	*Max pushed John. He fell*	FI^{ok}
(13)	Jean tomba. Max le poussa	*John fell. Max pushed him*	FI^{ok}, BI*
(14)	Max poussa Jean. Il était tombé	*Max pushed John. He was fallen*	BI^{ok},FI*
(15)	Jean tomba. Max l'avait poussé	*John fell. Max had pushed him*	BI^{ok}

Note here that the constraint imposed by tenses differs from the one imposed by a contextual assumption: no alternative reading than FI and BI are here possible. Optimal discourses will be those that violate the less constraints. Thus (12) and (15) are more optimal than (13) and (14).

---

[5] We assume that principles A and B are universal, whereas principle C is not. The distinction between incorporated and propositional information can be irrelevant in many languages.

3.3 Principle C

What happens when we add functional material like temporal or causal connectives? The prediction is that DI is given by connectives, which implies that connectives can impose a time direction opposite to the one encoded by tenses. If this is the case, the discourse should be interpretable, but less optimal than in cases where PI is co-directional:

(16) Max poussa Jean **et** il tomba	*Max pushed John and he fell*	FI^{ok}
(17) Jean tomba **et** Max le poussa	*John fell and Max pushed him*	FI^{ok}
(18) Max poussa Jean, **parce qu**'il était tombé	*Max pushed John, because he was fallen*	BI^{ok}
(19) Jean tomba, **parce que** Max l'avait poussé	*John fell, because Max had pushed him*	BI^{ok}

In these examples, where PI converge, (16) is more optimal than (17), and (19) more optimal than (18), because in (16) and (19), there are no conflicts between CI and PI, whereas in (17) and (18) such conflicts arise. In examples like (20) and (21) where PI diverges, interpretation processes become more difficult, and judgments of acceptability vary. For instance, (20) could be rejected from a normative point of view, whereas (21) is almost impossible to process because of the divergence between the time directions encoded by the connective (*et*) and by the tense (PQP):

(20) Jean tomba, parce que Max le poussa	*John fell, because Max pushed him*	BI^{ok}
(21) Max poussa Jean et il était tombé	*Max pushed John and he was fallen*	$FI^{??}$

We must now introduce two other principles explaning the way we combine in computation temporal and directional information. We have to introduce a new concept, that is, *directional feature*.

3.4 Directional features and principles D and E

Linguistic and non-linguistic information are hierarchical in the sense given by principles A to C. But the computation of DIs is based on atomic information encoded in directional *features*. We will examine here only two different features, forward and backward features. Features can on the other hand be weak or strong. So the DI device contains four types of features: (i) weak forward features [ff] ; (ii) weak backward features [bf] ; (iii) strong forward features [FF] ; (iv) strong backward features [BF].[6]

How can we then compute DIs relatively to directional features? We must introduce two subsequent principles:

D. A strong features wins over a weak feature or a string of weak features.
E. A weak feature or a string of weak features must be licensed by a strong feature.

These principles state that for a DI to be driven a strong feature must be accessible in order to give or to license the DI. The following algorithm makes explicit the procedure of DI assignation:

Algorithme of DI assignation

1. assign to utterance U1 a directional feature in function of directional features born by U1's linguistic expressions;
2. if possible, construct of a contextual assumption on the basis of conceptual expressions;
3. assign to utterance U2 a directional feature in function of directional features born by U2's linguistic expressions;
4. compute DI of discourse [U1-U2];
5. license DI *via* an accessible contextual assumption.

---

[6] The hierarchy of information given by principles A to C is not reflected in the typology of features: we are looking for general principles and do not want to have an *ad hoc* hierarchy of features

Here is an illustration of this algorithm, which gives a crucial role to contextual assumption in the licensing of DI:

Jean tomba, parce que Max l'avait poussé (John fell, because Max had pushed him)

1. assignation of directional feature to U1 (*Jean tomba*)
   tomba = $[ff_{PS}]$
   U1 : $[ff]_{U1}$
2. no contextual assumption can be formed from the concept *fall*
3. assignation of directional feature to U2 (*parce que Max l'avait poussé*)
   parce que = $[BF_{PQ}]$
   avait V-é = $[bf_{PQP}]$
   poussé = $[bf_{P}]$
   U2: $[BF_{PQ}]$ & $[bf_{PQP}]$ & $[bf_{P}]$ = $[BF]_{U2}$
4. computation of [U1-U2]'s DI
   [U1-U2]: $[ff]_{U1}$ & $[BF]_{U2}$ = $[BF]_{U1\text{-}U2}$
5. license DI *via* an accessible contextual assumption (CA):
   accessible CA: if Max pushes John, then John falls
   $[BF]_{U1\text{-}U2}$.

4. CONCLUSION

In this paper, I have provided an explicit model for the computation of DI in discourse. This model gives a crucial role to contextual assumptions, in the sense that they license or block the computed DI. Moreover, I provide principles that organise not only the different types of linguistic expressions, but also the way pragmatic computation works. Finally, the model of directional inference does not require any principle of discourse analysis, let alone principles of narrative discourse.

REFERENCE

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

Asher N. & Lascarides A. (2003), *Logics of Conversation*, Cambridge, CUP.

Blakemore D. (1987), *Semantic Constraints on Relevance*, Oxford, Basil Blackwell.

Caenepeel M. & Moens M. (1994), "Temporal structure and discourse structure", in Vet C. & Vetters C. (eds), *Tense and Aspect in Discourse*, Berlin, de Gruyter. 5-20.

Dowty D.R. (1986), "The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics?", *Linguistics and Philosophy* 9, 37-61.

Kamp H. & Rohrer C. (1983), "Tense in texts", in Bäuerle R. , Schwarze C. & von Stechow A. (eds), *Meaning, Use, and Interpretation of Language*, Berlin, de Gruyter, 250-269.

Labov W. & Waletzky J. (1967), "Narrative analysis: oral versions of personal experience", in Helm J. (ed.), *Essays on the Verbal and Visual Arts*, Seattle, University of Washington Press, 12-44.

Lascarides A. & Asher N. (1993), "Temporal interpretation, discourse relations and commonsense entailment", *Linguistics and Philosophy* 16, 437-493.

Moeschler J. (2000). "Le Modèle des Inférences Directionnelles ". *Cahiers de linguistique française* 22, 57-100.

Moeschler J. (2003), "Economy and pragmatic optimality: the case of directional inferences", *GG@G* 3, 1-20.

Reboul, A. & Moeschler J. (1998), *Pragmatique du discours*, Paris, Armand Colin.

Sperber D. & Wilson D. (1995), *Relevance. Communication and Cognition*, Oxford, Basil Blackwell.

Vet C. (1995), "Text type and discourse interpretation", in *Time, space and movement. Meaning and knowledge in the sensible world. TSM'95*, Toulouse, Université Toulouse-Le Mirail, 181-192.

Wilson D. and Sperber D. (1993a), "Linguistic form and relevance", *Lingua* 90/1-2, 1-25.

Wilson D. and Sperber D. (1993b), "Pragmatique et temps", *Langages* 112, 8-25.