

# ECONOMY AND PRAGMATIC OPTIMALITY: THE CASE OF DIRECTIONAL INFERENCES\*

Jacques Moeschler ([Jacques.Moeschler@lettres.unige.ch](mailto:Jacques.Moeschler@lettres.unige.ch))

## 1. INTRODUCTION

Economy has been one of the major topics in pragmatic theory, in two different meanings of the expression. First of all, in Gricean pragmatics the use of general and pragmatic principles (Cooperation Principle, maxims of conversation) tried to simplify linguistic explanation of utterance comprehension.<sup>1</sup> Roughly speaking, the function of pragmatic explanation was to allow a minimal (economical) semantic description, while explaining meaning in use as the result of conventional or conversational implicatures, depending on whether inference processes were triggered by the use or the exploitation of conversational maxims. Second, in the Gricean framework again, economy appears in a very broad sense of the word: an utterance respecting the maxims of quantity, quality, relation and manner was supposed to be more economical than an utterance, which would ostensibly violate them. Nevertheless, because of the loose and approximate definitions of these pragmatic principles, we still did not have a precise picture of how and why language use and utterance comprehension should obey any principle of economy.

Since Relevance Theory (RT)<sup>2</sup>, economy has mainly been used with much more precise functions: first to explain how cognitive processes are linked to utterance interpretation (processing efforts must be balanced by cognitive effects), second to explain how communication may be successful (inferences complete the underspecified content of the utterance to obtain its intended meaning). In other words, economy is a property of the cognitive systems devoted to utterance interpretation, as well as required, in order to insure a successful communication, by the computational devices, which combine with linguistic decoding to yield the intended meaning of the utterance.

In this contribution, even if I will not give a RT definition of the principle of economy, I would like to give some content to a notion close to economy, the notion of optimality.<sup>3</sup> Optimality is a new topic in pragmatics, and here again, RT is the first theory that tried to approach its cognitive and communicative functions. In RT, an utterance is relevant if and only if its cognitive effects balance the cognitive efforts its treatment requires. Apart from the definition of *relevance* (the more cognitive efforts, the less relevance; the more cognitive

---

\* This paper presents an up-dated version of the directional model first presented in an unpublished paper (Moeschler 1997), whose main proposals have been explicated in Moeschler et al. (1998, chapter 13) and Moeschler (1999a). A general presentation of the directional model is given in Moeschler (1998), Moeschler (1999b), Reboul & Moeschler (1998a, chapter 5) and Moeschler (2000). A previous version was read in Lyon at the CILISC1 (*Economy in structures, computations and use of language*, ISC, 12-15 October 1999) and at the ESF Exploratory Workshop in the field of the Humanities (*Temporal Reasoning in Discourse: Linguistic Variation and Cognitive Structure*, ISC, 23-25 February 2000). Thanks to Anne Reboul for her help and Laurence Danlos, Jean-Marie Pierrel, Bertrand Gaiffe, Dan Sperber for their comments.

<sup>1</sup> Cf. Gazdar (1979), Levinson (1983), Horn (1989).

<sup>2</sup> Cf. Sperber & Wilson (1986), (1995).

<sup>3</sup> A great deal of what will be say about optimality is close to propositions coming from Optimality Theory (OT) (cf. Archangeli & Langendoen (eds) 1997, Barbosa et al. (eds) 1998, Krager 1999).

effects, the more relevance), RT has proposed a definition of *optimal relevance*, linked to pragmatic expectations. The hearer is expecting a contribution that will modify his cognitive environment, but he has at the same time a presumption of optimal relevance: the utterance is relevant enough to be worth his while. So, minimal expectations meet a presumption of optimal relevance and it is in this way that RT gives sense to economy and optimality in utterance comprehension.

One question that RT has not yet answered in a theoretical way is how much linguistically encoded information contributes both to economy and optimality in utterance comprehension. One basic assumption of RT is that *utterance interpretation is linguistically underspecified*. This assumption will here be taken for granted, but I will question how linguistic information combines with non-linguistic information (that is, contextual information, encyclopaedic information) to achieve pragmatic interpretation. My basic assumption is that the combination of linguistic and non-linguistic information is directed by a general principle of optimality. This principle states that an optimal interpretation minimizes the conflict of information: the less conflicts you meet, the more optimal the interpretation you get.

This general assumption will be exemplified throughout a well-known problem in discourse semantics, the problem of *directional inferences*, that is, inferences concerning the temporal order (Narration) or the reverse order (Explanation) between events. The model I will present here, the *directional model* (DM), is a relevance-oriented framework devoted to solve the traditional problem of temporal order in discourse. I will not here present the classical (semantic and pragmatic) solutions to temporal order: I would like to show which kind of information are implied in the computation of directional inferences, and how natural language offers solutions to the cognitive systems implied in the computation of directional inferences. My basic assumption will be that relevance is achieved in the computation of directional inferences in as much as conflicts between directional properties borne by linguistic and non-linguistic information are minimal. The computation of directional inference implies a lot of information specialized in time direction. Thus the computational device must be able, under precise conditions, to license directional inference and to rule out other interpretations. The test for the more or less optimal linguistic solution will be increasing degrees of oddity in case of conflict between directional information borne by different types of information. Finally, I will raise the question of redundancy and parsimony in temporal discourse and question economy and optimality regarding the repetition or the lack of directional information encoded linguistically.

## 2. DIRECTIONAL INFERENCE

Let us begin by defining *directional inferences*. Directional inferences are inferences between utterances that indicate whether relations between eventualities have forward or backward temporal directions. Forward inference (FI) is illustrated in (1), whereas backward inference (BI) is the preferred reading for (2):

- |     |  |    |
|-----|--|----|
| (1) | Marie a poussé Jean. Il est tombé.<br>'Mary pushed John. He fell.' | FI |
| (2) | Jean est tombé. Marie l'a poussé.<br>'John fell. Mary pushed him.' | BI |

As the only formal difference (apart from pronominalization) is the linguistic order of sentences in discourse, how will the hearer make the right inference, that is, FI in (1) and BI in (2)? Before giving the core assumptions of the analysis, let us point out the following facts.

1. As far as time direction is concerned, BI is inferred in (1) and FI in (2) if connectives are added in those discourses:

(3) Marie a poussé Jean parce qu'il est tombé. BI  
'Mary pushed John because he fell.'

(4) Jean est tombé et Marie l'a poussé. FI  
'John fell and Mary pushed him.'

2. Time direction can be changed within a particular tense. Though, as shown by (1) and (2), the French Passé Composé allows both forward and backward readings, Passé Simple and Plus-Que-Parfait seem to allow only one reading (respectively forward and backward):

(5) a. Marie poussa Jean. Il tomba. FI  
'Mary pushed John. He fell.'

b. Jean tomba. Marie le poussa. FI  
'John fell. Mary pushed him.'

(6) a. Marie poussa Jean. Il était tombé. BI  
'Mary pushed John. He was fallen.'

b. Jean tomba. Marie l'avait poussé. BI  
'John fell. Mary had pushed him.'

3. If we combine observations given in 1 and 2, there seems to be a pragmatic constraint on time direction:

(7) a. Marie poussa Jean et il tomba. FI  
'Mary pushed John and he fell.'

b. Marie poussa Jean, parce qu'il tomba BI ?  
'Mary pushed John, because he fell.'

(8) a. Jean tomba et Marie le poussa. FI  
'John fell and Mary pushed him.'

b. Jean tomba, parce que Marie le poussa. BI  
'John fell, because Mary pushed him.'

(9) a. Marie poussa Jean et il était tombé. FI ?  
'Mary pushed John and he was fallen.'

b. Marie poussa Jean, parce qu'il était tombé. BI  
'Mary pushed John, because he was fallen.'

(10) a. Jean tomba et Marie l'avait poussé. FI ?  
'John fell and Mary had pushed him.'

b. Jean tomba, parce que Marie l'avait poussé. BI  
'John fell, because Mary had pushed him.'

A first conclusion can be drawn: all combinations of verbal tenses and connectives are not allowed. The results are not really incoherent, but some degree of oddity appears. We will

later on propose an explanation for this state of affairs. But let us resume briefly the empirical observations we can draw:

- (i) Examples (1) and (2) give the linguistic order of the sentence a crucial role. But this observation must be limited: only with unmarked past tenses like the French *Passé Composé* are FI and BI both possible. But we have to add another parameter: the order of the verbs in the sentence(s) considered give substantial indication for directional inference. We have thus the following preferred and dispreferred readings:

(11) Preferred readings (*passé composé*)

- a. #pousser# #tomber# → FI  
b. #tomber# #pousser# → BI

(12) Dispreferred readings (*passé composé*)

- a. #pousser# #tomber# → BI  
b. #tomber# #pousser# → FI

- (ii) The use of French past tenses in the second sentence, apart from the *Passé Composé* which is not directional, give univocal time direction: the *Passé Simple* impose FI, even if the preferred order of predicates is not compatible; the same holds with the *Plus-Que-Parfait*, specialized in BI. Thus, both preferred (11) and dispreferred (12) readings are ruled out by (13) and (14):

(13) Univocal interpretation with the *Passé Simple* (second sentence)

- a. #pousser# #tomber# → FI  
b. #tomber# #pousser# → FI

(14) Univocal interpretation with the *Plus-Que-Parfait* (second sentence)

- a. #pousser# #tomber# → BI  
b. #tomber# #pousser# → BI

- (iii) The use of connectives like *et* and *parce que* has a general effect: to force a particular directional inference, that is, FI with *et* and BI with *parce que*. These interpretations can contradict what we observed in (13) and (14), that is, the fact that French *Passé Simple* and *Plus-Que-Parfait* have univocal interpretations. So the rules associated to connectives like *et* and *parce que* seem to be the following:

- (15) a. *et* implies FI.  
b. *parce que* implies BI.

Now, we should draw some conclusions of these interpretations, and give a slightly different picture of the directional inference problem.

The first (provisory) conclusion is that information we have to take into account comes from three different sources: lexicon (verbs), verbal tenses and connectives. The first source of information belongs to a lexical category, defines an open class and makes accessible the representation of the eventuality (event, state). The second source of information is morphological: in French, tenses are affixes to verbs, belong to a closed class, and give information on how we have to treat the eventuality in the discourse. Finally, connectives belong to functional categories (conjunctions, adverbs), have sentences as scope and their main function is to connect representations of eventualities accessed and treated by the

combination of lexical and morphological categories. The question is now how combine these information.

The second conclusion can shed some light on this problem. One solution would be compositional. But as far as time direction is concerned, compositionality, even if required as we shall see, must be under control of a constraint on the strength of information. Clearly, and even if results vary on appropriateness and oddity, there seems to be a scale in the force of information : connectives are stronger than tenses (cf. (7)-(10)), and tenses are stronger than verbs (cf. (5)-(6)).

To make sense of this proposition, I must introduce a new distinction, first introduced in RT, the distinction between *conceptual* and *procedural* information.

### 3. CONCEPTUAL AND PROCEDURAL INFORMATION IN RELEVANCE THEORY

Following a proposition by Blakemore (1987:144), Wilson & Sperber (1993) made a very important distinction between two types of linguistically encoded information: conceptual information and procedural information.<sup>4</sup> Conceptual information is information about the concepts linguistic expressions represent, whereas procedural information is information about how to manipulate mental representations accessed by concepts.

What we meet here is something analogous. Our utterances are about eventualities, that is, roughly speaking events and states, and about relations between these eventualities. Now, suppose our cognitive system can access mental representations of eventualities, create new representations from old ones and new information. What we need is information on how to manipulate these representations, for instance erase an up-dated representation, order eventualities, include a specific one in a general one, etc. This is exactly what conceptual information and procedural information are supposed to do respectively.

Now imagine a symbolic system like a natural language, whose cognitive interface has access to both conceptual and procedural information. If natural languages are efficient, they should have specialized some of their expressions for these functions. My hypothesis is that lexical categories map onto conceptual information and functional categories map onto procedural information. The open/closed class opposition is thus cognitively motivated: we need to access a great amount of mental representations, and lexicon is a good tool to access concepts linked to these representations, but we do not want to have an infinite number of operations on these representations.<sup>5</sup>

We can now make sense of the hierarchy of information responsible for the computation of directional inferences, that is, of the hierarchy given in (16):

(16) connectives >> tenses >> verbs

Let us look first at a consequence of this hierarchy: it means that procedural information is stronger than conceptual information, as stated in (17):

(17) procedural information >> conceptual information

<sup>4</sup> In Blakemore (1987) (cf. also Blakemore 1992), the conceptual /procedural distinction belongs to the theory of meaning, that is, semantics: "On the one hand, there is the essentially *conceptual* theory that deals with the way in which elements of linguistic structure map onto concepts - that is, onto constituents of propositional representations that undergo computations. On the other, there is the essentially procedural theory that deals with the way in which elements of linguistics structure map directly onto computations themselves - that is, onto mental processes". Whereas I agree with her definitions, I give them a pragmatic content inasmuch as non-linguistic information participate to both conceptual and procedural information.

<sup>5</sup> In recent papers (Moeschler 2002 and to appear), I propose the conceptual/procedural opposition applies to any type of grammatical category. See Fraser (2001) for a similar analysis.

Why? The main argument is that information about representation of events or states cannot decide which operation you can apply to these representations. It thus seems reasonable to distinguish conceptual information from procedural information, and to let instructions on how to manipulate representations dominate information in those representations.

We would now justify the first part of the hierarchy in (16), that is, the hypothesis that connectives are stronger than tenses. The argument is based on the domain of these functional categories. While tenses are affixes and compose with stems inflected verbal forms, connectives are conjunctions or adverbs having sentences as scope. The reason is then easily found: the greater the scope, the stronger the directional force. (18) makes sense of this hierarchy:

(18) phrasal domain >> lexical domain

We have now the beginning of a description, based on a pragmatic framework devoted to the computation of directional inferences. Before introducing the core notion of the directional model, that is, *directional features*, I need to make a small comment on examples, which seem to be counterexamples to the proposed hierarchy.

#### 4. APPARENT COUNTEREXAMPLES

If we look at examples (7b), (9a) and (10a), predictions on the strength hierarchy are ruled out:

- |      |    |   |      |
|------|----|---|------|
| (7)  | b. | Marie poussa Jean, parce qu'il tomba.<br>'Mary pushed John, because he fell.' | BI ? |
| (9)  | a. | Marie poussa Jean et il était tombé.<br>'Mary pushed John and he was fallen.' | FI ? |
| (10) | a. | Jean tomba et Marie l'avait poussé.<br>'John fell and Mary had pushed him.'   | FI ? |

First note that both connectives *et* and *parce que* are concerned here. The classical answer would be, indeed, that if *parce que* could be a strong marker of BI, *et* would be too weak to force FI. This is partially right, because the argument we gave against the strong directional property of tenses can be applied to connectives: all uses of *parce que* do not imply BI, as well as all uses of *et* do not imply FI.<sup>6</sup>

How can we now explain the bad results in these discourses? Let us examine quickly those examples.

(i) In example (7b), the reading should be BI, because *parce que* is stronger than the Passé Simple and the order of lexical item <*poussa-tomba*> is irrelevant here. But the BI reading is very strange, and if a reading is possible and accessible, it is the FI one. This is a very curious fact, but if we look at *parce que* in non-temporal discourse, we observe the same phenomenon:

<sup>6</sup> Cf. Luscher (1994), (1999) and Moeschler (1989) for detailed analyses of *parce que* and *et*.

- (19) Il y a du poulet froid dans le frigo, parce que je n'ai pas envie de faire à manger.  
'There is cold chicken in the fridge, because I do not want to cook.'

In this example, a causal BI does not capture the speaker's informative intention. What she means is that the presence of chicken in the fridge allows one to conclude that she does not want to cook. So the FI in (7b) is not an exception. But (19) is generally considered odd. The explanation is that a forward information as strong as the backward information of *parce que* is accessible. Here, the causal relation between *pousser* and *tomber* allows one to build a contextual forward assumption, that is, an assumption allowing one to order the events parallel to the discourse order. This contextual assumption is the origin of possible anticipatory hypotheses, like *Jean will fall*.

(ii) The two other odd examples have the same properties with reverse order of predicates and tenses: the second clause in the Plus-Que-Parfait follows a Passé Simple and *et* explicates the connection. It is this combination of tenses and connective, which makes the connection strange and difficult to understand. (9a) is consistent relative to order of events and connection, but inconsistent relative to the Plus-Que-Parfait, and (10b) exhibits the opposite properties. There seems to be here major conflicts of directions. But the question is why this conflict is much more serious than in (8a) and (9b) for instance. To answer this question implies introducing the core notion of the directional model, that is, *directional features*.

## 5. DIRECTIONAL FEATURES

What we need now is a way of capturing directional properties of linguistic expressions and contextual assumptions. The simplest proposal is to associate to each kind of information (conceptual, procedural, and contextual) directional features. In principle, we have a three-ranked hierarchy, given in (16):

- (16) connectives >> tenses >> verbs

But we know, while discussing (7b), that contextual information is the strongest one, because it forces a FI reading. So we get a four-ranked hierarchy, given in (20):

- (20) contextual assumptions >> connectives >> tenses >> verbs

This scale exhibits another scale, defining the type of the information:

- (21) contextual information >> procedural information >> conceptual information

Finally, (21) can be reduced to the opposition contextual/linguistic information:

- (22) contextual information >> linguistic information

Now we see why a quantitative approach to directional strength of information is not appropriate: it would neither capture the nature of the information bearer, nor the grammatical hierarchy. We need thus a qualitative definition of time direction, and an evaluation the source of the information. We make thus the following proposal:

- (23) a. Verbs and tenses bear weak directional features.  
b. Connectives and contextual assumptions bear strong directional features.

These features (strong or weak) have a direction. I write strong features in capitals and weak features in standard fonts, *f* for forward feature and *b* for backward feature, and give the following conventional direction to directional expressions:

- (24) a. Tenses : Passé Simple = *f*, Plus-Que-Parfait = *b*  
 b. Connective : *et* = *F*, *parce que* = *B*  
 c. Verbs : *pousser* = *f* in first position, *b* in second position<sup>7</sup>

As we can infer time direction without any strong features (if no connective is explicated for instance), we have to give the key on how directional features combine, relative to their intrinsic direction, their force and their source. If we look precisely at the combination of time directions for our basic examples, as tableau 1 shows, we can distinguish different types ('→' means 'accommodation of directional features'):

Examples	Verbs	Tenses	Connectives	Contextual assumptions	Interpretations
5a ☞ <i>Marie poussa Jean.</i> <i>Il tomba.</i>	<i>f</i>	<i>f</i> <i>f</i>		<i>F</i>	<i>FI</i>
5b <i>Jean tomba.</i> <i>Marie le poussa.</i>	<i>b</i>	<i>f</i> <i>f</i>		<i>B</i> ® <i>F</i>	<i>FI</i>
6a <i>Marie poussa Jean.</i> <i>Il était tombé.</i>	<i>f</i>	<i>f</i> <i>b</i>		<i>F</i> ® <i>B</i>	<i>BI</i>
6b ☞ <i>Jean tomba.</i> <i>Marie l'avait poussé.</i>	<i>b</i>	<i>f</i> <i>b</i>		<i>B</i>	<i>BI</i>
7a ☞ <i>Marie poussa Jean</i> <i>et il tomba.</i>	<i>f</i>	<i>f</i> <i>f</i>	<i>F</i>	<i>F</i>	<i>FI</i>
7b <i>Marie poussa Jean,</i> <i>parce qu'il tomba.</i>	<i>f</i>	<i>f</i> <i>f</i>	<i>B</i>	<i>F</i>	?? <i>FI</i>
8a <i>Jean tomba</i> <i>et Marie le poussa.</i>	<i>b</i>	<i>f</i> <i>f</i>	<i>F</i>	<i>B</i> ® <i>F</i>	<i>FI</i>
8b <i>Jean tomba,</i> <i>parce que Marie le poussa.</i>	<i>b</i>	<i>f</i> <i>f</i>	<i>B</i>	<i>B</i>	<i>BI</i>
9a <i>Marie poussa Jean</i> <i>et il était tombé.</i>	<i>f</i>	<i>f</i> <i>b</i>	<i>F</i>	<i>F</i>	? <i>FI</i>
9b <i>Marie poussa Jean,</i> <i>parce qu'il était tombé.</i>	<i>f</i>	<i>f</i> <i>b</i>	<i>B</i>	<i>F</i> ® <i>B</i>	<i>BI</i>
10a <i>Jean tomba</i> <i>et Marie l'avait poussé.</i>	<i>b</i>	<i>f</i> <i>b</i>	<i>F</i>	<i>B</i> ® <i>F</i>	?? <i>FI</i>
10b ☞ <i>Jean tomba, parce que Marie</i> <i>l'avait poussé.</i>	<i>b</i>	<i>f</i> <i>b</i>	<i>B</i>	<i>B</i>	<i>BI</i>

Tableau 1: a first analysis of directional inferences

- Two kinds of situation do not raise problems, either because all features have the same direction, as in (5a) and (7a), or because a converse weak feature in the first tense is ruled out by a string of backward features, as in (6b) and (10b). They are mentioned by the sign '☞' meaning they are optimal.

<sup>7</sup> A direction is given to contextual assumptions on the basis of either encyclopaedic information or accessible contextual information.

2. The second set of examples requires a change in time direction of contextual assumption. In (5b), (6a), (8a) and (9b), a weak verbal feature give rise to a cancelled expectation and must be erased in favour of the opposite direction. In each case, the order of event is *fall* < *push*.
3. A third group of examples illustrates various judgments of acceptability due to the choice of the verbal tense in the second clause (8b) or of the connective (7b). In the first case, the tense features are *f*, and all other features are backward. Here, the consistency depends on norms. A group of French native speaker rejects (8b) because the only French possible tense is the Plus-Que-Parfait, as in (10b). (7b) is generally rejected because of the contradiction of the B feature (*parce que*) contradictory with all weak and strong forward features. Generally, French native speakers evaluate this discourse as odd if not unacceptable.
4. Finally, (9a) and (10b) are very strange because of the interpretation of the second clause containing a Plus-Que-Parfait, which should imply a BI: the presence of *et* impose a FI reading, and make the compatibility of *F* (*et*) and *b* (Plus -Que-Parfait) difficult, if not impossible.

What conclusion can we draw from this description? We can first give a more precise definition of what an optimal configuration in discourse is: *an optimal configuration in discourse minimizes conflicts between directional features*. When a potential conflict arises, pragmatic accommodation is the only way to save appearances. This is possible if accessible or constructible situations correspond to directional inferences triggered linguistically. If conflicts arise, either between weak and strong features or between strong features, pragmatic accommodation is more or less acceptable depending on linguistic or discourse norms.

These conclusions could be interpreted in two different ways: in the first one, no generality can be drawn, because of the idiosyncratic nature of pragmatic accommodation; in the second one, generalizations are possible, but these rules can be defeated. I will defend the second solution, because the first one gives no explanation neither to the regularity of linguistic properties nor to the generality of inferential processes used by participants in communication.

Here is a set of rules of the directional model:

- I. *A string of co-directional weak features must be confirmed by a strong feature to give time direction*. Examples (5a) is the prototype of this situation: its forward reading is possible not only because of weak conceptual and procedural features *f* borne by verbs and tenses, but because a contextual hypothesis with *F* feature is accessible from conceptual information.
- II. *A string of non-co-directional weak features is solved by a strong feature*. This situation is illustrated by example (6b) and (10b), where a B feature resolves the conflict between *f* and *b* features.
- III. *A linguistically encoded strong feature can be deleted if a contextual assumption requires it*. This is the case with the exceptional change of directional feature of a connective (7b).
- IV. *Pragmatic accommodation arises when a first contextual assumption is reversed because of evidences coming from the world (accommodation goes from encoded information to contextual information)*. This is the case with examples (8a), (9b) and (10a), with various degrees of acceptability.
- V. *An optimal interpretation minimizes directional conflicts*. Optimal discourse are here (5a), (6b), (7a) and (10b).

This is, I think, a positive conclusion. In spite of the diversity of linguistic and contextual resources, there seems to be a general character governing discourse production as

well as discourse interpretation: the minimization of informational conflict between directional features. Optimality is thus what completes economy, because pragmatic accommodation is in some case the only path to get the intended interpretation.<sup>8</sup>

## 6. PRAGMATICS AND LINGUISTICS

I tried to show what the relations between linguistically encoded information and contextual information are. I made a crucial hypothesis, which I will repeat here:

*Contextual information is stronger than linguistic information.*

What are the empirical evidences that support this assumption? The main argument is that when directional conflicts occur, accommodation is generally pragmatic. In other words we accommodate the directional inference relatively to the relevant accessible contextual assumption.

This hypothesis is nothing new in pragmatics, if we admit the thesis of underspecified linguistic interpretation. But we have now to answer to a new question:

- (i) Where is the economy of such a device if linguistically encoded information can be cancelled? This question must be completed by another one:
- (ii) Why can linguistically encoded directional information be repeated in discourse?

Let us try answering the first question. The question of cancellation of directional features is linked to the question of their strength. Why do we assign to verbal tenses weak features and not strong features? Why do we assign a directional feature to the logical particle *et* (and)? To answer these questions is, partly at least, to answer the questions of redundancy and parsimony of directional information.

Until now, I gave no account of the structural properties of sentences. The reason is obvious: I choose very simple sentences and very simple discourse connections, assuming that no structural properties of sentences would affect the computation of directional features. But let now take our first example (5a), repeated here, and imagine a way of representing structurally directional features:

- (5) a. Marie poussa Jean. Il tomba.  
'Mary pushed John. He fell.'

Here are the structural descriptions of these two sentences.

---

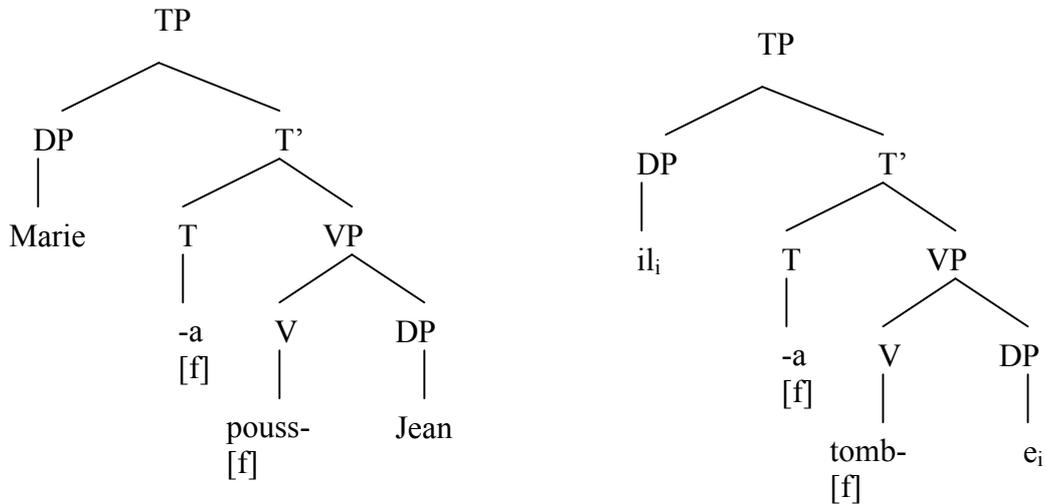
<sup>8</sup> There seems to be gaps in language about procedural expressions. For instance, in French, there is no forward causal connective, that is, a connective that could do the job of the following instruction:

- a. *P conn Q*
- b. interpret *P* as the cause of *Q*.

Some uses of *parce que* do this job, but no other item can: *donc* is not causal, but conclusive. One way speakers can fill the gap of no causal forward connective is to use *parce que* in a front position, as in (c) with (d) interpretation:

- c. *parce que P, Q*
- d. Interpret *P* as the cause of *Q*.

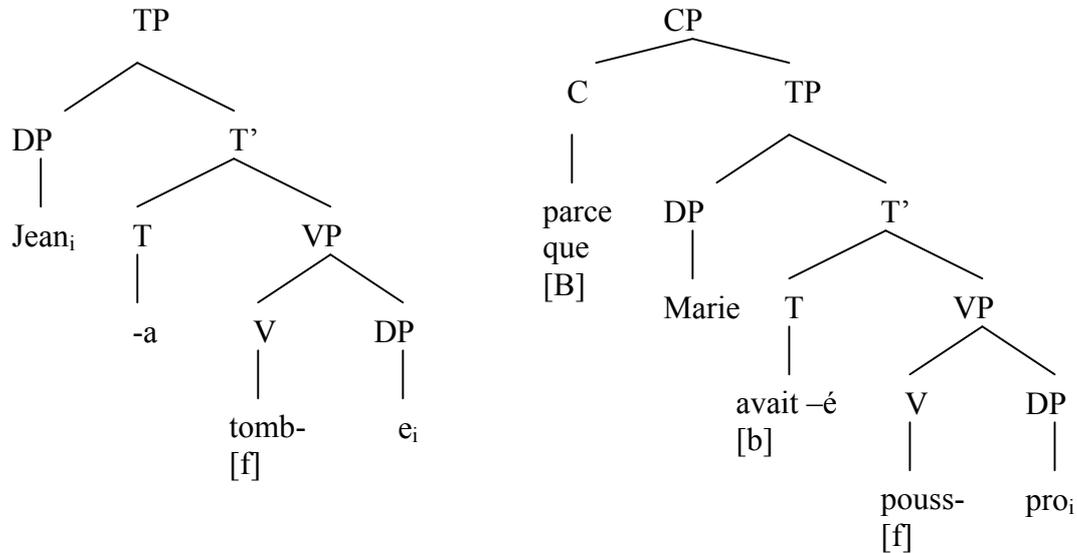
Another strategy is to use, as (19), *parce que* inferentially. See Moeschler (1989) for further arguments.



What these trees show, in a very simplified manner, is that directional features are assigned to different categories (T and V) and after incorporation an inflected lexical item (*poussa*) contains two directional features. So what the structural representation shows here is that directional information is not redundant, because it is associated to different categories. Redundancy is thus something superficial. Thus what is important is that directional information can be propagated along the structural representation so that it can emerge as a final interpretation.

Suppose now we have to parse (10b), which is optimal in terms of directional information, with only one conflict in tense direction, solved by the other backward features. We can have the following syntactic representations (cf. Berthouzoz 2000 for further analyses from syntax to pragmatics):

- (10) b. Jean tomba, parce que Marie l'avait poussé.  
'John fell, because Mary had pushed him.'



The presence of three directional features in the second clause insures its time direction. No redundancy is present here, because of the different categorial nature of the time direction bearer. Hence we have an optimal way of transporting directional information.

Suppose now we have (6b) instead of (10b), that is no causal connective.

- (6) b. Jean tomba. Marie l'avait poussé.  
 'John fell. Mary had pushed him.'

The syntactic representation is simpler, because no lexical item occupies the C position and bears a strong directional feature. But there is no interpretive difference: the BI reading is as accessible in (6b) as in (10b). Moreover, (10b) requires more cognitive effort, and should thus be less relevant than (6b) regarding directional inference.

This answer is not completely correct, because *parce que* brings another information than directional information: *parce que* makes the causal relation between events explicit. In other words, the causal relation is implicated in (6b) whereas it is explicated in (10b). The cancellation of causal relations is possible without the presence of *parce que*, but impossible with *parce que*:

- (25) a. Jean tomba. Marie l'avait poussé. Jean avait bien réussi à se raccrocher à la barrière, mais le coup porté par Marie le mit définitivement à terre.  
 'John fell. Mary had pushed him. John had succeeded in catching the gate but Mary's kick finally made him fall.'  
 b. ?? Jean tomba, parce que Marie l'avait poussé. Jean avait bien réussi à se raccrocher à la barrière, mais le coup porté par Marie le mit définitivement à terre.  
 'John fell, because Mary had pushed him. John had succeeded in catching the gate but Mary's kick finally made him fall.'

If the causal relation is implicated in (6b), it is because a causal relation can be inferred from conceptual information (a *push-fall* causal rule). So the previous answer is not complete. Even if there is a difference between an explicated causal relation and a implicated causal relation, which one is the most optimal? In other words, when a speaker has to communicate directional and causal relations, which (by implicature or by explicature) is the best way to achieve this relation? We have two possible answers:

1. The first answer says that explicating relations is a way of avoiding ambiguity or misunderstanding. In this way, redundancy insures communication.
2. The second type of answer says that explicating causal relation is a mean of ensuring intended meaning when no evidence can be drawn from linguistically encoded information. In this way, there is no redundancy, and explicit connection is necessary for a successful communication.

The second type of answer is much more interesting than the first one. The hypothesis is that the presence or absence of some procedural expressions like connectives depends on the nature of other types of information. Mainly, if conceptual information does not allow to draw the correct inference, a connective is required. As tenses are obligatory within a finite verb, the false choice gives rise to a wrong interpretation, as in (5b) with a BI intended meaning:

- (5) b. Jean tomba. Mary le poussa.  
       ‘John fell. Mary pushed him.’

What are the conclusions of this discussion? Mainly that the repetition of directional information is not a matter of ensuring communication, but a question of accessibility of contextual information: when contextual information can be drawn from conceptual information and no pragmatic accommodation is necessary, no higher level bearer of directional information is required. So economy has to be defined as a balance between accessible contextual information and linguistically encoded information.

But we can draw another conclusion, concerning the relation between linguistics (restricted to phonology, syntax, semantics) and pragmatics. Pragmatics is not any more a wastebasket, that is, a bunch of information that linguists do not know how to represent and figure out. Pragmatics interacts with linguistics in a very subtle way, because language processing in communication implies taking into account information about the world. This interrelation between linguistic and pragmatic information is the key to an understanding of the way economy and optimality work in language use.

Economy requires that no unnecessary cognitive effort should be spent in utterance interpretation. Presumption of optimal relevance requires thus that cognitive effects will balance the efforts engaged in utterance comprehension.

We saw that optimality is a way of making sense of this proposition. An optimal way to communicate directional inference will be not only to minimize the cognitive costs, but also to minimize conflicts between directional features. We can thus formulate two principles for interpreting directional inferences:

*Principle of economy: minimize cognitive efforts.*

*Principle of optimality: minimize conflicts in directional features.*

Following these two principles, we can propose a measure of optimality for our test examples. Here is the following optimality scale:

Groups		Examples	Verbs	Tenses	Connectives	Contextual assumptions	Interpretations
A	5a	<i>Marie poussa Jean.</i> ☞ <i>Il tomba.</i>	<i>f</i>	<i>f</i> <i>f</i>		<i>F</i>	<i>FI</i>
	6b	<i>Jean tomba.</i> ☞ <i>Marie l'avait poussé.</i>	<i>b</i>	<i>f</i> <i>b</i>		<i>B</i>	<i>BI</i>
	7a	<i>Marie poussa Jean</i> ☞ <i>et il tomba.</i>	<i>f</i>	<i>f</i> <i>f</i>	<i>F</i>	<i>F</i>	<i>FI</i>
	10b	<i>Jean tomba,</i> ☞ <i>parce que Marie l'avait poussé.</i>	<i>b</i>	<i>f</i> <i>b</i>	<i>B</i>	<i>B</i>	<i>BI</i>
B	8b	<i>Jean tomba,</i> <i>parce que Marie le poussa.</i>	<i>b</i>	<i>f</i> <i>f</i>	<i>B</i>	<i>B</i>	<i>BI</i>
C	8a	<i>Jean tomba</i> <i>et Marie le poussa.</i>	<i>b</i>	<i>f</i> <i>f</i>	<i>F</i>	<i>B</i> ® <i>F</i>	<i>FI</i>
	5b	<i>Jean tomba.</i> <i>Marie le poussa.</i>	<i>b</i>	<i>f</i> <i>f</i>		<i>B</i> ® <i>F</i>	<i>FI</i>
	6a	<i>Marie poussa Jean.</i> <i>Il était tombé.</i>	<i>f</i>	<i>f</i> <i>b</i>		<i>F</i> ® <i>B</i>	<i>BI</i>
	9b	<i>Marie poussa Jean,</i> <i>parce qu'il était tombé.</i>	<i>f</i>	<i>f</i> <i>b</i>	<i>B</i>	<i>F</i> ® <i>B</i>	<i>BI</i>
D	9a	<i>Marie poussa Jean</i> <i>et il était tombé.</i>	<i>f</i>	<i>f</i> <i>b</i>	<i>F</i>	<i>F</i>	? <i>FI</i>
	7b	<i>Marie poussa Jean,</i> <i>parce qu'il tomba.</i>	<i>f</i>	<i>f</i> <i>f</i>	<i>B</i>	<i>F</i>	?? <i>FI</i>
	10a	<i>Jean tomba</i> <i>et Marie l'avait poussé.</i>	<i>b</i>	<i>f</i> <i>b</i>	<i>F</i>	<i>B</i> ® <i>F</i>	?? <i>FI</i>

Tableau 2: Scale of optimality in directional inferences

Let us make a quick comment on this scale. Like our precedent analysis, four groups can be distinguished:

- A. In the first group (5a, 6b, 7a, 10b), the difference is between the explicit or the implicit character of a strong directional feature. The principle of economy states that the implicit way of expressing directional inference is more optimal than the explicit way, because a strong directional feature can be inferred from conceptual (lexical) information.
- B. In the second group, containing only one example (8b), the conflict is between weak procedural features (tenses) and all other types of features. But no accommodation is required, because no strong feature has to be changed.
- C. The third group contains examples illustrating pragmatic accommodation (8a, 5b, 6a, 9b). Here accommodation implies the change of time direction inferred by conceptual information, either because tenses and connectives have the opposite direction to conceptual information, or because tenses only are opposite to lexicon.
- D. Finally, in the last group (9a, 7b, 10a), there is a conflict between verbs, tenses and connectives, and all interpretations are difficult to access.

## 7. CAVEATS AND CONSEQUENCES

I tried, in this paper, to give a fine grained analysis of the directional inferences possible from discourse generated with two predicates, two tenses and two connectives. This is a typical fragment of French. The limits of my proposals give rise to two questions:

- (i) How can we say that this model is general enough to capture directional inferences in discourse implying different predicates, different verbal tenses, different connectives, and more than two sentences?
- (ii) What kind of relation between semantics and pragmatics does the directional model imply?

Let us begin with question (i). The first point to note is that directional inferences are not the only possible temporal relations between eventualities. We saw that in Backward Inference (BI), causality is also the case. Discourse semantics, and more precisely Asher's SDRT<sup>9</sup>, give a precise definition of other types of discourse relations than Narration (FI) and Explanation (BI). For temporal discourses, Elaboration, Background and Results are also relevant.

A. *Results* imply causality and forward inference, but are generally associated to resulting states, described in French with the Imparfait, as in (26):

- (26) a. Jean alluma la lampe.  
'John switched the lamp on.'
- b. La faible lumière donnait à la pièce un air de tristesse.  
'The weak light was giving the room a sad look.'

The description of the Imparfait must be compatible with the FI reading. Clearly, as discussion on uses of Imparfait shows, the resulting interpretation is due to temporal implications inferred from the first sentence. So a strong contextual feature is inferred from the conceptual relation between *allumer* and *lumière*. The neutral directional property of the imparfait is thus ruled out by the forward direction inferred by contextual information.

B. *Elaboration* defines a part-whole relation, which can be captured in terms of temporal inclusion, as (27) shows it:

- (27) a. Guy passa une merveilleuse soirée hier soir.  
'Guy experienced a lovely evening last night.'
- b. Il fit un repas extraordinaire.  
'He had a fantastic meal.'

This example, taken from Lascarides & Asher (1993), is very interesting, because we have to block FI (Narration) between *a* and *b*, in favor of a part-whole relation (*b* is a part of *a*). Clearly, a model of temporal relation must add temporal inclusion as result of temporal inference. In a pragmatic model of temporal relations, these relations are inferred from conceptual information. Thus, the following contextual assumption should be accessible from conceptual information:

- (28) A meal is a part of an evening.

---

<sup>9</sup> Asher (1993), Lascarides & Asher (1993), Asher et al. (1995).

C. Finally, *Background* implies temporal *inclusion* and *unboundness*, as illustrated in (29):

- (29) Jean entra. Marie téléphonait.  
 ‘John entered the room. Mary was on the phone.’

In order to get the inclusion relation, we have to complete the directional model and add two other temporal features: an *inclusion feature* (*in*), a *bounded* (*bo*) and an *unbounded* (*ubo*) *feature*.<sup>10</sup> In Result, both forward and unbounded features must emerge. In Elaboration, inclusion must be inferred from part-whole relation between eventualities. Finally, in Background, inclusion and unboundness are the case.

The second point I would mention is that the directional model must be able to make computations on more than two sentences. The main problem for a discourse theory is to be able to attach new information to old information given in previous utterances. In standard discourse theory (like SDRT), the principle is to give explicit conditions on discourse relations and to create a new SDRS including the relevant discourse referents for the connection. Another solution (Saussure 2000) is to attach new information to temporal variables and connect events to each other via their temporal variables.

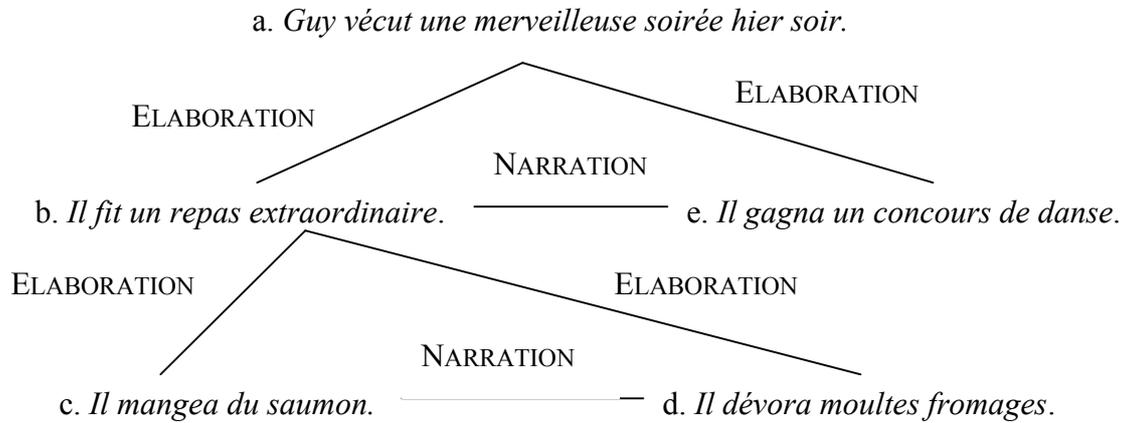
The solution implied by the directional model is a little bit different. I will not develop it, but give the flavor of the solution. When you have to treat a complex discourse, several solutions are offered to attach a new eventuality to the other ones. Suppose that two information are conveyed linguistically, an *in* feature lexically encoded (by predicates) and an *f* feature functionally encoded (by tenses). This situation is illustrated by the complete example given by Lascarides & Asher (1993):

- (30) a. Guy passa une merveilleuse soirée hier soir.  
 ‘Guy experienced a lovely evening last night.’  
 b. Il fit un repas extraordinaire.  
 ‘He had a fantastic meal.’  
 c. Il mangea du saumon.  
 ‘He ate salmon.’  
 d. Il dévora moult fromages.  
 ‘He devoured lots of cheese.’  
 e. Il gagna un concours de danse.  
 ‘He won a dancing competition.’<sup>11</sup>

This is a very interesting example, because a stupid parser would give false results: *FI* with local attachment. This is inappropriate, because the discourse structure, as proposed by Lascarides & Asher (1993), is the following:

<sup>10</sup> The French Imparfait has inclusion and unbounded features.

<sup>11</sup> I give the French version because of the explicit presence of the Passé Simple, associated with *f*.



In other words, there are part-whole relations (Elaboration) between *a-b*, *a-e*, *b-c* and *b-d*, and FI (Narration) between *b-e* and *c-d*. The problem is to get the right attachments.

What the Directional Model allows us to do is to license temporal inclusion when a strong *IN* feature is derived from a weak lexical feature. If we assume that the following contextual assumptions are accessible, then inclusive interpretations between *a-b*, *b-c*, *b-d* and *a-e* are accessible:

- (31) a. A meal is a part of an evening.  
 b. Fish is a part of a meal.  
 c. Cheese is a part of a meal.  
 d. A dance competition is a part of an evening.

How to explain then FI between *c-d* and *b-e*? In our model, we must complete the weak feature *f* assigned to the Passé Simple by a strong contextual feature *F*. The answer is to use an analogous principle given in SDRT, in which two elaboration relations having a same node as source give rise to Narration between targets of the relation.<sup>12</sup> So from *IN(b, c)* and *IN(b, d)*, we infer *FI(c, d)*. The same hold for *FI(b, e)* from *IN(a, b)* and *IN(a, e)* (cf. Tableau 3):

Discourse	Verbs	Tenses	Contextual assumptions	Interpretations
a <i>Guy vécut une merveilleuse soirée hier soir.</i>		<i>f</i>		
b <i>Il fit un repas extraordinaire.</i>	<i>in</i>	<i>f</i>	<i>IN</i>	<i>IN(a, b)</i>
c <i>Il mangea du saumon.</i>	<i>in</i>	<i>f</i>	<i>IN</i>	<i>IN(b, c)</i>
d <i>Il dévora moult fromages.</i>	<i>in</i>	<i>f</i>	<i>IN</i> <i>F</i>	<i>IN(b, d)</i> <i>F(c, d)</i>
e <i>Il gagna un concours de danse.</i>	<i>in</i>	<i>f</i>	<i>IN</i> <i>F</i>	<i>IN(a, e)</i> <i>F(b, e)</i>

Tableau 3: a directional analysis of a small complex discourse

So the computation of complex discourse relations is possible within the directional model.

Nevertheless, I have to answer the second question concerning the relation between truth-functional semantics and pragmatics within the directional model. What I said previously about the relation between linguistics and pragmatics concerns mainly the relation

<sup>12</sup> This rule is ruled out in case of non-ordered encapsulation (cf. Saussure 2000).

between syntax and pragmatics: what the linguistic parser gives to pragmatics is not only structural relations between constituents, but also directional features allowing inferential computation on time direction. What is lacking is a semantic interface, capable of assigning truth conditions to clauses.

This traditional picture, which assigns to semantics the role to capture truth-functional properties of sentences, will be amended here. My proposal is based on two assumptions coming respectively from discourse semantics and RT.

The first assumption is that the truth-functional property of a sentence is relative not only to the eventuality it describes but also to the type of relation between eventualities. For instance, the difference between bounded (*bo*) and unbounded (*ubo*) eventualities determines the nature of temporal relations entertained (inclusive inference *INI* or forward inference *FI*), as shown in the two following discourses:

- |         |  |     |
|---------|--|-----|
| (32) a. | Jean entra. Marie téléphonait [ <i>ubo</i> ]<br>'John entered the room. Mary was phoning.'   | INI |
| b.      | Jean entra. Marie téléphona [ <i>bo</i> ]<br>'John entered the room. Mary was on the phone.' | FI  |

The consequence is that the truth-functional properties of clauses depend on the computation of directional inferences. So what classical analysis calls truth-functional semantics occurs after pragmatics.<sup>13</sup>

The second assumption comes from RT and the distinction between logical form (LF) and propositional form (PF). LF is a string of concepts, the output of the linguistic input systems. The combination of contextual information and LF produces FP, which is an enrichment of LF. It is at PF that disambiguation, referent assignment, as well as aspectual class determination occur. So the truth-functional level is no more LF, but PF. In other words, semantics is not enough to get a complete truth-functional interpretation, but pragmatics is.<sup>14</sup>

Within these two assumptions, we can give a much more precise picture on how linguistics interacts with pragmatics:

- I. The linguistic device produces structural representations, including directional and temporal features.
- II. LF is derived and contains conceptual and procedural information, that is, addresses for concepts and directional features for procedural information.
- III. At the pragmatic stage, selection or derivation of contextual assumptions is made, and computations of directional features take into account linguistic and contextual information.
- IV. Once directional information is computed, PF and truth-value can be assigned.

We can try to give the following general picture, with the illustration of *parce que Marie l'avait poussé*:

<sup>13</sup> See Levinson (2000) for further arguments.

<sup>14</sup> Cf. Moeschler & Reboul (1994) and Reboul & Moeschler (1998b) for further developments.



propositional forms. The truth-functional interpretation of the utterance is thus a pragmatic and not a semantic process.

## REFERENCES

- Archangeli, D. & D.T. Langendoen (eds) (1997), *Optimality Theory. An overview*, Basil Blackwell, Oxford.
- Asher, N. (1993), *Reference to Abstract Objects in Discourse*, Kluwer, Dordrecht.
- Asher, N. *et al.* (1995), “De l’espace-temps dans l’analyse du discours”, *Sémiotiques* 9, 11-62.
- Barbosa, P. *et al.* (eds) (1998), *Is the Best Good Enough? Optimality and Competition in Syntax*, MIT Press, Cambridge, Mass.
- Berthouzoz, C. (2000), “Le Modèle Directionnel d’Interprétation du Discours”, *Cahiers de Linguistique Française* 22, University of Geneva, 101-146.
- Blakemore, D. (1987), *Semantic Constraints on Relevance*, Basil Blackwell, Oxford.
- Blakemore, D. (1992), *Understanding Utterances. An Introduction to Pragmatics*, Basil Blackwell, Oxford.
- Fraser, B. (1991), “On the conceptual-procedural distinction”, ms. Boston University.
- Gazdar, G. (1979), *Pragmatics. Implicature, Presupposition, and Logical Form*, Academic Press, New York.
- Horn, L.H. (1989), *A Natural History of Negation*, The University of Chicago Press, Chicago.
- Krager, R. (1999), *Optimality Theory*, Cambridge University Press, Cambridge.
- Lascarides, A. & N. Asher N. (1993), “Temporal interpretation, discourse relations and commonsense entailment”, *Linguistics and Philosophy* 16, 437-493.
- Levinson, S.C. (1983), *Pragmatics*, Cambridge University Press, Cambridge.
- Levison, S.C. (2000), *Presumptive Meanings. A Theory of Generalized Conversational Implicatures*, MIT Press, Cambridge, Mass.
- Luscher, J.M. (1994), “Les marques de connexion: des guides pour l’interprétation”, in J. Moeschler *et al.*, *Langage et pertinence. Référence temporelle, anaphore, connecteurs et métaphore*, Presses Universitaires de Nancy, Nancy, 175-227.
- Luscher, J.M. (1998), *Éléments d’une pragmatique procédurale. Le rôle des marques linguistiques dans l’interprétation*, Ph.D., University of Geneva, Department of linguistics.
- Moeschler, J. (1989), *Modélisation du dialogue. Représentation de l’inférence argumentative*, Hermès, Paris.
- Moeschler, J. (1997), “Inférences en avant et inférences en arrière”, ms., University of Geneva, Department of linguistics.
- Moeschler, J. (1998), “Directional inferences and the conceptual/procedural encoding distinction”, *Proceedings of the Relevance Theory Workshop*, University of Luton (8-10 september 1998), 3-8.
- Moeschler, J. (1999a), “Linguistique et pragmatique cognitive. L’exemple de la référence temporelle”, *Le Gré des Langues* 15, 10-33.
- Moeschler, J. (1999b), “Representing events in language and discourse”, in M. Coene, W. de Mulder, P. Dendale & Y. D’Hulst (eds), *Hommage à Liliane Tasmowski-de Ryck*, Unipress, Padova, 461-479.
- Moeschler, J. (2000), “Le Modèle des Inférences Directionnelles”, *Cahiers de Linguistique Française* 22, University of Geneva, 57-100.
- Moeschler, J. (2002), “Connecteurs, encodage conceptuel et encodage procédural”, *Cahiers de Linguistique Française* 24, University of Geneva.

- Moeschler, J. (to appear), "Pragmatics and linguistic encoding. Evidence from the conceptual/procedural distinction", in P. Schulz (ed.), *Semiotics and Communication Sciences*, University of Toronto Press, Toronto.
- Moeschler, J. *et al.* (1998), *Le temps des événements. Pragmatique de la référence temporelle*, Kimé, Paris.
- Moeschler, J. & A. Reboul (1994), *Dictionnaire encyclopédique de pragmatique*, Le Seuil, Paris.
- Reboul, A. & J. Moeschler (1998a), *Pragmatique du discours. De l'interprétation de l'énoncé à l'interprétation du discours*, Armand Colin, Paris.
- Reboul, A. & J. Moeschler (1998b), *La pragmatique aujourd'hui. Une nouvelle science de la communication*, Le Seuil (Points), Paris.
- Saussure, L. de (2000), *Pragmatique temporelle des énoncés négatifs*, Ph.D., University of Geneva, Department of linguistics.
- Sperber, D. & D. Wilson (1986), *Relevance. Communication and Cognition*, Basil Blackwell, Oxford.
- Sperber, D. & D. Wilson D. (1995), *Relevance. Communication and Cognition*, Basil Blackwell, Oxford, 2<sup>nd</sup> edition.
- Wilson, D. & Sperber, D. (1993), "Linguistic form and relevance", *Lingua* 90/1-2, 1-25.