Constraints in Discourse 2

EDITED BY

Peter Kühnlein

Anton Benz

Candace L. Sidner

p & b

Constraints in Discourse 2

Edited by

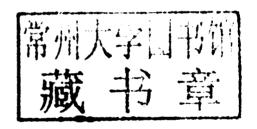## Peter Kühnlein
function2form

Anton Benz
Centre for General Linguistics, Berlin

Candace L. Sidner
Worcester Polytechnic Institute

Constraints in Discourse 2

Pragmatics & Beyond New Series (P&BNS)

Pragmatics & Beyond New Series is a continuation of Pragmatics & Beyond and its Companion Series. The New Series offers a selection of high quality work covering the full richness of Pragmatics as an interdisciplinary field, within language sciences.

Volume 194

Constraints in Discourse 2
Edited by Peter Kühnlein, Anton Benz and Candace L. Sidner

Table of contents

Rhetorical structure

An introduction

Peter Kühnlein

0.1 General remarks

Texts, and in general types of discourse, vary along a multitude of dimensions. Discourse can be spoken or written, monological or an exchange between a number of participants, it can be employed to inform, persuade (and serve many more or even mixed functions), it can take place in various settings and be arbitrarily extensive. However, some characteristics are shared by all kinds of texts.

One of those shared properties is that text, and discourse in general, is structured, and they in turn are so in a multitude of ways: classical written text as the present, e.g., typically has logical and graphical structuring into paragraphs, sections, chapters etc. depending on the type of text or the genre, but it is by nature monological. Spoken dialogue, at the other end of the spectrum, *inter alia* is characterized by assignment of and changes in roles participants assume in the exchange, stretches of overlapping speech, repairs and many more phenomena that are not regularly observed in written text (notwithstanding chats on the internet and the like) and which give rise to completely different types of structure. All of these forms of communication fall under the common denominator *discourse*; we will keep using this cover term here to refer to them.

The different kinds of structures in discourse have been object to research for a considerable time. One type of structure has been of special interests for researchers working in more formal paradigms and has been hotly discussed ever since: it is what is called the *rhetorical* or *coherence structure*. Rhetorical structure is built by applying rhetorical relations recursively to elementary units of discourse. This kind of structure is to be distinguished from, e.g., *cohesive* structure that comes to existence by means of, e.g., coreference in various forms.

The present collection comprises papers that give a wide variety of perspectives on the constraints governing discourse structure, and primarily rhetorical structure: various ways of thinking of constitutive units of discourse along with a variety of conceptions of rhetorical relations are presented, and the issue of which kind of structure is right for the description of the rhetorical make-up of discourse is tackled from different points of view.

Accordingly, this introductory chapter is intended to provide the necessary background to understand the discussions by sketching as briefly as possible the state of the art; the reader is referred to the individual chapters in the volume where appropriate.

As the previous volume, *Constraints in Discourse* (Benz & Kühnlein, 2008), the present one is the result of selecting and compiling papers that are extended versions of presentations at a workshop in the series "Constraints in Discourse." The second of these workshops was held in Maynooth, Ireland, and organized by Candace Sidner (chair), Anton Benz, John Harpur and Peter Kühnlein. All the authors who contribute to the present volume submitted their re-worked and substantially extended papers to a peer reviewing process, where each author had to review two other authors' papers. In addition, John Benjamins conducted an own reviewing process before agreeing to publish the collection. This two-stage reviewing process is intended to secure high quality of the contributions.

0.2   Elementary units

Just as in any formal description of structures, one basic step in describing rhetorical structure of a discourse is to identify the elementary units. Due to the multitude of dimensions along which discourse can vary and due to differences in theoretical assumptions, there is no consensus on what to count as an elementary unit.

Exemplarily, there is a divide between proposals for different domains: a proposal for spoken discourse can refer to intonational features as an important criterion for segment status, whereas a proposal made specifically for written discourse can't. On the other hand, a proposal set up for written discourse can make reference to punctuation and syntactic units, whereas the first is absent and the the latter are not reliably correct in spoken discourse.

Research in prosodic features of discourse and its segments reaches far back: Butterworth (1975) reports that speech rate changes during discourse segments, being higher at the end of a segment than at the beginning. Chafe (1980) observes that pause lengths are varying at segment boundaries too. Much corpus based and computer linguistic research in this area was conducted by Julia Hirschberg with various collaborators, e.g., Hirschberg and Pierrehumbert (1986), Grosz and Hirschberg (1992), Hirschberg and Pierrehumbert (1992), or Hirschberg and Nakatani (1996).

One of the most detailed empirical inquiries into the relation between discourse segmentation and prosody is given in (Hirschberg & Nakatani, 1996),

and the methodology employed there deserves a little closer description. The authors set up a corpus of directives, where subjects had to give route descriptions of varying complexity through Boston. The first series of descriptions was given spontanously by the subjects, recorded and then transcribed. In a second series, the same subjects read the corrected (i.e., freed of false starts etc.) transcripts of their route descriptions, and again the speech was recorded and transcribed. The data obtained from one of the speakers were prosodically transcribed using the ToBI standard, split into INTERMEDIATE PHRASES, pause lengths were measured and fundamental frequencies (F0) and energy (RMS) calculated.

Then two groups of annotators marked up the texts with segment boundaries: one group was given the transcripts only, the other group was given transcripts plus recorded speech. The theory that served as background to segmenting the texts was that of Grosz and Sidner (1986). Their account is potentially independent from domain, i.e., applicable to both spoken and written discourse; Grosz and Sidner claim that discourse structure actually consists of three distinct, but interacting levels. The most central of these levels is the intentional one: for every coherent discourse, that is the claim, one can identify an overarching discourse purpose the initiating participant seeks to pursue. The segments of discourse according to this theory correspond to sub-purposes, the so-called discourse segment purposes. Elementary units in this theory correspond to single purposes. The two other levels, attention and linguistic realization, concern which objects are in the center of discourse and how the discourse is actually realized using cue-phrases and special markers.

The results obtained by (Hirschberg & Nakatani, 1996) in their study on spoken discourse confirm previous findings and reveal much more detail than, e.g., the work by Butterworth (1975); Chafe (1980) reports: both F0 and energy are higher at the beginnings of discourse segments than at their ends. Speech rate on the other hand increases towards the end of a segment, and pause lengths during a segment are shorter than before a segment beginning and after a segment end. So segment boundaries as judged according to purposes indeed seem to be correlated with measurable changes in the speech signal.

These, and similar, findings seem to indicate good mutual support between the intention-based theory of discourse structure developed by Grosz and Sidner (1986) and the claim that discourse segment boundaries are marked intonationally in spoken discourse.

Considerably more work than on spoken discourse has traditionally been devoted to written text than to spoken discourse phenomena. The pioneering work dates back to the 80s, and the cited work by Grosz and Sidner (1986) is among that. An account of discourse structure that had comparable impact at that time was developed by Mann and Thompson (1987a, 1988b). This account,

known as Rhetorical Structure Theory (RST) was explicitly developed as a means to capture analysts' judgements about writers' intentions while composing texts. Thus, RST is devoted to the analysis of written text, but the analysis is not primarily guided by linguistic surface structure: according to its founders, it is rather "pre-realizational" in that it aims to describe the function of (the interplay of) constituents in abstraction from linguistic realization.

Mann and Thompson (1987a, 1988b) claim that the base case for linguistic expressions conveying intentions are clauses of certain types: main clauses, non-restrictive relative clauses are of the right variety, whereas, e.g., restrictive relative clauses and complement clauses (e.g., in subject or object position in a matrix clause) are not counted as minimal units. It turned out in the development of RST since its inception that narrowing down the type of constructions that express writers' intentions to clauses poses problems in multi-lingual applications: what is expressed in a clause in one language might more suitably be expressed in a different construction in another language. This case was made especially by Rösner and Stede (1992) and Carlson and Marc (2001) who consequently proposed extensions of the set of minimal units. The motivation for the inclusion of certain constructions (or the exclusion of others) is not always readily understandable. So, researchers comprising Carlson and Marcu (2001) and Lüngen et al. (2006) (cf. also paper 0.5) working in the RST paradigm, but likewise Wolf and Gibson (2005, 2006), opt for including certain pre-posed PPs like "On Monday," in the list of minimal units, whereas temporal adverbs that potentially convey the same information ("Yesterday,") are not included. One reason for this decision might be that those researchers are working on corpora based on journal texts, where temporal expressions preferably are of the variety they include in the list; so the decision to include one type of expression but not the other might be rather pragmatic than theory-driven.

Other work has been less intention-oriented than that of Grosz and Sidner and that in the RST paradigm, and consequently employed a different reasoning to select units as elementary discourse units. One line of research that also dates back to the mid-80s of the last century seeks to understand coherence in more general terms than tied up with linguistics. In (Hobbs, 1985) and work that can be seen in its tradition, like (Kehler, 2002), it is argued that coherence in text is by and large a product of the capability of rational agents to understand the world as being coherent. In this tradition, in its roots at least dating back to Hume (cf. (MacCormack & Calkins, 1913)), what is related by the rational mind are events or states of affairs. Consequently, what counts as a minimal unit in these accounts are expressions that can serve to convey states or events, or in short *eventualities*. A first class citizen here is the clause again, and once again with suitable restrictions excluding, e.g., restrictive relative clauses. For different reasons, Asher and Lascarides (2003)

and Reese et al. (2006, 2007) working in the Segmented Discourse Representation Theory (SDRT) paradigm too consider the expression of eventualities to be the decisive criterion for individuating minimal units.

As is well known from work in Montague grammar and elsewhere, it is all too easy to coerce the type of expressions to that of an eventuality. In fact, the paper by Jerry Hobbs in this collection (see 0.5) focuses on going below the clause level as minimal units. Hobbs there points out that he does think that even single words potentially express eventualities. It seems there is a thin line between raising types to that of an eventuality too easily and missing out on sub-clausal constituents that in fact are rhetorically interesting.

Yet another line of research different from the intention-oriented and the eventuality-based ones can be seen in processing-based accounts. One of the first candidates there, and also rooted in the mid-1980s is the work of Polanyi (1986, 1988); but also the work by Webber (2004) can be seen in that tradition. Polanyi's LDM most closely mirrors the incremental nature of text processing in that a discourse tree is built by adding sentences as elementary units to the existing representation of the text so far perceived. Sentences obviously are larger units than clauses since they potentially consist of multiple clauses (matrix, relative clauses, complement clauses). Both the LDM and Webber's D-LTAG suggest extensions to sentential syntax to model discourse structure, and thus there seems to be no need for sub-sentential segmentation.

Both LDM and D-LTAG — at least as concerns segmentation — thus seem to follow the opposite strategy than that pursued by Hobbs in his contribution in the present volume and reserve the rhetorical importance to larger units.

As a summary to the above approaches to segmentation, it seems that all accounts agree on a core set of units (main clauses that form sentences) that should be treated as elementary units, whereas there is large disagreement as to what else should be considered a unit in discourse.

0.3   Rhetorical relations

In Section 0.2 various views on how to split up discourse were reported. The present section is concerned with putting Humpty Dumpty together again: it is agreed among linguists that coherent discourse should be represented as a connected structure where each segment is connected to the rest by rhetorical relations. Islands in the representation of the analysis of a text are dispreferred and viewed as either a sign of incoherence of the discourse under analysis, faulty analysis itself or some lack in descriptive power in the inventory of rhetorical relations.

In what follows in this section, the accounts used to introduce segmentation strategies in the order chosen in Section 0.2 will be taken up in turn again and the rhetorical relations employed by those accounts will be sketched.

On the intention-oriented side, Grosz and Sidner (1986) employ a surprisingly small set of rhetorical relations. In their seminal paper they mention only two of them, one being *dominance* (the dominated discourse unit serves to achieve the goal of the super-ordinate) and the other *satisfaction precedence* (the preceding goal has to be achieved before the next can be achieved). The authors are aware of the fact that in, e.g., the work by Mann & Thompson (ultimately published in Mann and Thompson (1987a), but circulating in various grey versions beforehand) a much larger number of rhetorical relations are discussed. However, since for Grosz and Sidner primacy is on *intentions* and their relations to each other rather than on textual realizations of intentions, they can claim that dominance and satisfaction-precedence are sufficient for the description of rhetorical structure and the specific relations between textual units derivable from structure and intention content.

Grosz and Sidner (1986) explicitly set up their account for construction dialogues; given the goal orientation of that dialogue type, it seems that the inventory consisting of dominance and satisfaction-precedence suffices to describe the intentional structure of dialogues from that domain. This might be questioned, however, in a more general domain, where a putative task structure (if there is any) might be not as tightly bound to discourse structure. On the other hand, it has to be said that Grosz and Sidner (1986) do not *deny* the existence of more relations between discourse purposes. The claim, it seems, is just that the set suffices for the analysis of the given type of discourse.

As mentioned, Mann and Thompson (1987a, 1988b) posit a much larger set of discourse relations that should be used to describe the functional role of elementary units as recognized by the analyst: according to that classification, one unit might, e.g., *elaborate* on another, or units might form a *list*. There are two main divides within the class of relations: the first divide concerns the functional classification of relations: part of them are *subject-matter* relations (reporting about facts), another part is *presentational*, employed to influence the readers' stance towards the (main or local) discourse topic. Both of these types of relations can be realized in either of two ways (giving the second divide) — connecting a less important part of discourse (a *satellite*) to a more important one (called *nucleus*), or connecting nuclei to nuclei. The first type along the latter divide is called *mono-nuclear* relation (or nucleus-satellite relation), the second *multi-nuclear*.

One of the tests for nuclearity of discourse units is an elimination test: eliminating nuclei from a text tends to render it incoherent, while eliminating satellites

tends to leave coherence intact. This property of nuclei has led Marcu (1996) to posit the *nuclearity principle* for RST, claiming that spans of texts are connected by a relation iff their nuclei are. (A consequence of that principle for the representation of discourse structure will be discussed in 0.4.) According to Matthiessen and Thompson (1988), there is another indicator for nuclearity: they observed that there is a high correlation between the status of being a nucleus in a text and of being realized in a main clause just in case a rhetorical relation is present between two syntactically related clauses. Syntactically subordinate clauses tend to realize satellites in turn. As Matthiessen and Thompson (1988) warn, this is not a hard and fast rule, but rather a tendency, and counter examples abound. There even seem to be language specific discourse connectives that trigger an inversion in nuclearity, like the dutch connective *zodat* (*so that*) which in a majority of cases syntactically subordinates a nucleus to a satellite.

Bateman and Rondhuis (1994, 1997), and recently Stede (2008), systematically investigate rhetorical relations across different discourse theories and, for RST's nuclearity, propose not to tie the assignment of nuclearity to the presence of certain rhetorical relations (i.e., to drop the divide between mono-nuclear and multi-nuclear relations) and to view assignment of nuclearity as an effect of the presence of other factors. This seems to be in line with the findings by Asher and Vieu (2005) who claim something similar for an analogous divide among relations in SDRT. The insight that certain relations can be viewed as connecting nuclei to satellites *or* satellites to other satellites seems also to be the rationale behind the explosion of number of relations in the RST-flavor proposed by Carlson and Marcu (2001), where multi-nuclear versions of relations formerly categorized as mono-nuclear abound.

The discussion about the "right" relations for RST doesn't seem to be settled nor does it seem it has to be: Taboada and Mann (2006a) in their recent overview over developments in RST propose that researchers in the paradigm should tailor their own relations according to their specific needs for specific purposes.

The situation is different in SDRT, which, as mentioned, knows a similar divide as the nucleus-satellite distinction in RST. SDRT knows thorough axiomatizations of the discourse relations that are employed. These relations take the semantic representations of the minimal units and join them in either of two ways: by coordinating a unit to a preceding one, or by subordinating one unit to another. The nature of the relation involved (subordinating or coordinating) has influence on the possibilities where subsequent units can be attached: if the last relation involved was coordinating, then the constituent to which the last unit was related by it is blocked for attachment. If the last relation, on the other hand, was subordinating, then both the last unit and the one to which it was attached are available. These constraints on attachment points for new discourse units give rise to what is

called the *Right Frontier Constraint* (RFC), first postulated by (Polanyi, 1986). One effect of the RFC is to limit anaphoric accessibility: antecedents are said to be only available for (pronominal) anaphoric uptake if they occur in a unit that is on the right frontier.

Asher and Vieu (2005) re-examine the distinction between the two classes of relations and suggest that the question whether, e.g., a *cause*-relation is subordinating or coordinating depends in part on the information structure exhibited by the units that are related. Certain information structural configurations in the units can lead to anaphoric accessibility of discourse referents whereas truth semantically equivalent variants of that information structure makes them inaccessible. This fact can be accounted for if it is assumed that the information structure at least in part can change the way a unit is attached to preceding discourse.

SDRT draws another distinction between discourse relations that resembles the distinction between presentational and subject-matter relations in RST: many of the discourse relations are *content-level* relations which are similar to subject-matter relations, whereas other relations bear more resemblance to presentational relations, like *text-structuring, cognitive-level* discourse relations or *metatalk* relations. Interestingly, the so-called satisfaction scheme for veridical relations holds for some, but not all, relations of either variety. The latter scheme tells that two (representations of) discourse units connected by a relation are part of the interpretation of a discourse just in case the interpretations of the units are and the interpretation of the relation is. Whereas a veridicality criterion like that is to be expected for content-level relations, it is not so clear that a relation like *parallel* (a text-structuring relation) should have that property. None of the cognitive-level relations are veridical, though.

Just like SDRT and the account of Grosz and Sidner (1986), but unlike RST, the account of discourse relations given by Hobbs (1985); Hobbs et al. (1993) and, more or less based on it, Kehler (2002) is an attempt to give a principled way to define the ways units of discourse are combined to form larger units. Hobbs et al. (1993) distinguishes four classes of discourse relations: some that are inferred to hold because the units that are connected are about events in the world (like casual relations), others that relate what was said to an overall goal of the discourse, again others that relate a unit to the recipient's prior knowledge (e.g., *background*) and finally "expansion" relations (like *contrast*).

0.4   Structures and their properties

So, both what counts as minimal units and in which ways they can be combined by rhetorical relations are matters of dispute in discourse theory. Given this situation,

it can be expected that there is also no consensus on which structures discourse can be expected to have.

The expectation is confirmed by the literature. Whereas many researchers — e.g., Polanyi (1986, 1988, 2001), Grosz and Sidner (1986, 1998), Mann and Thompson (1987b, 1988a), Taboada and Mann (2006b) — assume that trees suffice to model the rhetorical structure of dialogue, there is an increasing number of theorists that doubt this assumption for a variety of reasons.

Prominent among the latter are Wolf and Gibson (2005, 2006) who argue for a much less constrained type of graphs for the description of rhetorical structure. The *chain graphs* they postulate as adequate for the description of rhetorical structure feature all kinds of violations of tree structure: they posit nodes with multiple parents, crossing edges, and in general graphs without a distinguished root node. Their strongest constraint on structures seems to be connectedness and acyclicity. These graphs are capable of describing all kinds of relations between elementary units; a closer look at their annotation manual and the set of relations they employ reveals that this seeming strength is a real weakness too: the set of relations Wolf and Gibson (2005, 2006) employ is a mixed bag, mostly taken from (Hobbs, 1985; Hobbs et al. 1993), but considerably modified and enriched with some relations from (Carlson & Marcu, 2001). The annotation manual requires analysts to annotate not only rhetorical relations used for combining minimal units, but also coreference relations and other cohesive devices that can be present within minimal units as well. Their first step of analysis, *grouping*, actually consists in connecting units that are related by cohesive links. Only after that very step rhetorical relations are applied to the units — alas not to the units *connected* by the first step. Thus, it is no wonder that crossing dependencies and nodes with multiple parents abound in the analyses presented in Wolf and Gibson (2005, 2006). Knott (2007) questions the statistics Wolf and Gibson (2005, 2006) perform on their data, claiming that the small percentage of tree violations that can be tied to the special relations introduced by Wolf and Gibson in their evaluation of their data is implausible. I don't think so: rather, Knott's critique seems to set in too late. The true reason for the high amount of tree violations does not lie in the special relations, but in the conflation of levels of analysis.

Another line of attack on tree structures as the adequate description for rhetorical structures can be found in (Danlos, 2004, 2008). Danlos compares the generative capacity of a comparatively unrestricted formalism (an extension of Mel'čuk's (Roberge, 1979) dependency syntax to discourse) with those of RST and SDRT. She derives all the structures that can be generated by either formalism and tries to find discourses that exhibit the respective structure. Her benchmark formalism generates directed acyclic graphs (DAGs), whereas RST and SDRT generate trees. According to Danlos analysis, RST undergenerates (is not complete), her

benchmark formalism overgenerates (is not correct), and SDRT is closest to being both complete and correct, with the exception being a few structures that can not be described as SDRT-trees. Danlos argument to my mind has two drawbacks, though it is admirably ingenious. First, it should be evaluated on corpus data instead of relying on constructed discourse for confirmation. This is mainly a precaution against an overreliance on intuitions, of course. The second point is a bit stronger: the reconstruction of RST mainly — based on (Marcu, 1996) and (Carlson & Marcu, 2001) — seems to contain too strong an interpretation of the nuclearity principle that leads to the assumption of graphs that are not in accordance with most other work in RST. So, rather than raising an argument against general RST assumptions, her attack is directed against a very idiosyncratic version of RST.

But both the account of Wolf and Gibson (2005, 2006) and of Danlos (2004, 2008) are under active discussion, and until there is conclusive evidence to the contrary, it has to be assumed that there are strong arguments against general tree-hood of discourse structure. A weaker warning comes from Webber (2001) and Lee et al. (2008): these authors caution that although most discourse can be mod-elled as trees, there might be certain cases where a departure from tree structures is required. So the warning would be to give up the general claim in favour of a rule of thumb with dened exceptions.

It seems that the question of how rich a structure has to be assumed for the description of discourse these days is more hottly debated than ever. The papers in the present volume will help to solve or focus in this debate by contributing insights in the fundamental questions that have to be answered.

0.5   About the papers

Jerry R. Hobbs: Clause-Internal Coherence

As was discussed in Section 0.2, there is no unanimity about the size or gen-eral characterization of elementary discourse units, just as there is no agreement on the definitions of relations between them. In his contribution to this volume, Hobbs extends his account, e.g., from (Hobbs, 1985), to cover coherence at a sub-clausal level.

Henk Zeevat: Optimal Interpretation for Rhetorical Relations

Zeevat argues that rhetorical relations can be reconstructed from general optimality theoretic (OT) assumptions. He gives a comprehensive introduction to OT, with special emphasis on the constraints *NEW, RELEVANCE, FAITH and PLAUSI-BLE. He then continues to demonstrate how a range of rhetorical relations can be derived from a certain ordering of these constraints; most importantly, *NEW and

RELEVANCE tend to introduce rhetorical structure defaults, with PLAUSIBLE being a filter over the generated relations. This account of coherence relations is in marked contrast to accounts such as that of Hobbs (1985) or Asher and Lascarides (2003).

Ekatarina Jasinskaja: Modelling Discourse Relations by Topics and Implicatures: The Elaboration default

Jasinskaja argues in her paper for the position that discourse relations can be inferred by utilising underlying pragmatic principles such as *topic continuation* and *exhaustive interpretation* as defaults. In the absence of linguistic markers that make one more inclined to infer a different relation, she opts for ELABORATION as one of the default relations, since it best obeyes both principles, i.e., does not induce topic shifts and at the same time add information to the topic at hand.

Maja Bärenfänger, Harald Lüngen, Mirco Hilbert & Henning Lobin: The role of logical and generic document structure in discourse analysis

The authors of this contribution propose to add two descriptive levels to the local rhetorical analysis of discourse structure: the logical structure (like title, paragraph etc.) and the genre specific structure (introduction, method). Structure at these levels is usually not explicitly signalled, yet conventionalized, and can thus be used to guide (automatic) parsing of texts. The authors strive to clarify which cues and constraints can be observed at these levels of discourse and demonstrate their utility for automatic text processing.

Pascal Amsili & Claire Beyssade: Obligatory Presupposition in Discourse

Presupposition triggers have been considered obligatory under certain conditions by a variety of authors. One of the conditions that was deemed necessary in previous work was that the triggers are additive particles, like *too*. Amsili & Beyssade argue that this condition is not a necessary one, but that obligatoriness is the case for triggers that have no asserted content. (*Too* being but one of them.) They give a general explanation for the apparent sensitivity of this class of triggers to discourse relations and provide a formalization in terms of an SDRT update mechanism, building on Asher and Lascarides (1998).

Ann Copestake & Marina Terkourafi: Conventionalized speech act formulae — from corpus findings to formalization

Copestake & Terkourafi present an account to the semantics and pragmatics of conventionalized speech acts which renders the contribution of the illocutionary force as an addition to the compositional semantics of the utterance. They motivate their account with examples from a corpus of Cypriot Greek and formalize it within the framework of HPSG. They show how their account leaves the possibility for a literal interpretation of conventionalized speech act formulae open, thus opening the possibility to react to them in a variety of ways.