LU

LINGUISTISCHE
UNTERSUCHUNGEN

Gerd Fritz

Coherence in Discourse

A Study in Dynamic Text Theory

Giessen University Library Publications 2022

Coherence in Discourse. A Study in Dynamic Text Theory

Linguistische Untersuchungen 16

Herausgegeben von Gerd Fritz, Thomas Gloning und Dennis Kaltwasser

LU

LINGUISTISCHE
UNTERSUCHUNGEN

Gerd Fritz

Coherence in Discourse

A Study in Dynamic Text Theory

Giessen University Library Publications 2022

Schlagwörter

Textkohärenz, Dialogkohärenz, Sprechaktsequenzen, Kontextveränderung, Themenentwicklung, Textdynamik, Dialogdynamik

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie. Detaillierte bibliografische Daten sind im Internet unter http://dnb.d-nb.de abrufbar.

ISBN: 978-3-347-70940-9
DOI: http://dx.doi.org/10.22029/jlupub-791

Giessen University Library Publications, Gießen 2022

Verlag und Druck: tredition GmbH, Halenreie 40-44, 22359 Hamburg
Umschlaggestaltung: Harald Schätzlein · ultraviolett.de

Contents

Preface

Variants of the concept of coherence – or rather different concepts of coherence – play a role in various academic fields, from the theory of truth to cognitive psychology, the social sciences, and linguistics. Generally speaking, coherence seems to be a property of the way we humans see the world, projecting forms of connectedness and consistency onto given phenomena, and of the way we act in various contexts. In his recent book on coherence, Givón even views an orientation towards coherence as "a property of biologically-based systems" in general (Givón 2020, 1). Now, the notion of discourse coherence that I shall deal with in the present book is much more restricted and is based on the everyday intuition that texts and (contributions to) conversations mostly show a kind of orderliness and serial connectedness that makes them comprehensible and useful. Lack of coherence, on the other hand, sometimes makes communication difficult or even impossible. What is probably also part of our relevant everyday intuitions is the fact that there are degrees of coherence/incoherence and that it is sometimes hard to say what exactly it is that makes a discourse "hang together". It is sometimes much easier to pinpoint cases of incoherence, which happen accidentally, e.g., in cases of misunderstanding, or are produced deliberately, e.g., in modern literature, where the intentional disruption of text structures can contribute to a distancing effect and other aesthetic effects, and in comedy, where incoherence is used to achieve a comical effect. Dreams – and dream narratives – may also appear incoherent. Children's speech sometimes sounds incoherent to grown-ups, and the speech of persons with aphasia and schizophrenia has also often been described as incoherent. And it is maybe not surprising that texts produced by Artificial Intelligence models have been known to contain incoherencies. Thus, questions of coherence have complex ramifications.

As for my own interest in discourse coherence, it was in the early 1970s that I first became aware of a new field of research called text linguistics and a basic question asked in this field: What makes texts coherent? Reading research in this field I soon felt that an approach that basically followed structuralist principles (e.g., Harweg 1968, Halliday/Hasan 1976) was unsatisfactory and that ideas and methods being developed in pragmatics might help to create a more satisfactory approach to problems of coherence. This motivated me to do research in this area, the results of which were published in a book with the title "Kohärenz. Grundlagen der linguistischen Kommunikationsanalyse" (1982). (So, writing the present book was for me a case of "coherence revisited"). In writing the earlier book I was inspired by Wittgenstein's concept of language games and research developing this idea (and related ideas),

e.g., Hamblin's dialectics with its concept of "commitment store", early versions of context-change theories, Alston's version of a speech-act theory, Strawson's (and others') theories of reference, Grice's "Logic and conversation", and also ethnomethodological studies on sequencing phenomena, among others.

In my later research I found that problems of coherence arose in various contexts, including the study of comprehensibility and usability, historical text types, structures of dialogue, forms of controversy, science communication in digital media, multimodal discourse, and the structure of hypertext. And, of course, monitoring my own writing processes and marking students' papers were useful sources of information on coherence. A recent addition to my interests consisted in the methods and results of computational discourse processing. This interest surfaces in my discussion of the achievements of GPT-3 (in chapter 11).

The experiences gained in this work also influenced my theoretical thinking on matters of discourse and coherence. As a kind of summary of this theoretical thinking and its possible applications I wrote my "Dynamische Texttheorie" (2017), in which questions of coherence naturally played an important background role and also explicitly came to the fore in many contexts, but were not presented as a topic in its own right. Exploring this topic in more detail is what the present book aims to do. What may be considered new in this book is (i) the emphasis on the *interaction* of resources applied in the production of coherent discourse, e.g., common ground, topics, and speech act sequencing patterns, (ii) the emphasis on the *dynamics* of the use of these resources in common ground management, context change, topic management, and speech act sequencing, and (iii) the combined treatment of monologue and dialogue from the same theoretical perspective. In some passages I will be discussing competing theories, e.g., Rhetorical Structure Theory and Centering Theory, however my intention is more constructive than polemical.

In working on the projects mentioned above I enjoyed the cooperation and creativity of many colleagues and friends, from whom I learned a lot that is relevant to my present work. I cannot mention them all here. (Some of them are mentioned in relevant passages of the book.) I have, however, in writing this book particularly benefitted from the knowledge and patience of a number of friends who read (large parts of) the manuscript and offered thoughtful reflections on the text, in particular Hans-Jürgen Bucher, Dennis Kaltwasser, and Thomas Gloning, whose helpful comments and constructive suggestions I gratefully acknowledge. Finally, I would like to thank my wife Renate for many inspiring discussions on the topics of this book and for her patience and encouragement, and, last but not least, my daughter Fiona for her interest in this work and her kind help with technical problems.

1. Introduction

What is discourse coherence? The everyday view of coherence seems to be that a coherent text or argument is one in which the parts somehow "hang together" and are not inconsistent. Such a text is comprehensible and makes sense. In order to come to grips with this somewhat vague concept, linguists and others have to search for the factors and aspects of discourse that make it appear well-organised and connected.

In early linguistic work on coherence, it was assumed that coherence was (mainly) a surface property of texts that is based on the presence of certain "cohesive" elements connecting a sequence of sentences or clauses. As one of the objections against this view it was noted that a given sequence of sentences could be considered either coherent or incoherent, depending on the context assumed for an interpretation of this sequence. This observation, again, lead some researchers to define coherence as a "context-dependent, hearer- (or reader-) oriented and comprehension-based *interpretive* notion" (Bublitz 1999, 2; cf. Stede 2007, 188). A similar view is taken by psychologists and psycholinguists who see coherence as a property of the mental models or representations of a text which hearer/readers construct in the process of comprehension (e.g., Kintsch 1998, 307ff.; Gernsbacher 1997).

This view, however, only captures part of the role of coherence in communication, and it is remarkable that authors who spend much of their time and effort on *producing* coherent texts themselves should take such a one-sided view. The other side of the coin is the fact that speaker/writers *create* a coherent text or contribution to conversation by using various textual and contextual resources. Therefore, discourse coherence is also and arguably even primarily a notion related to the performance of communicative actions and the solution of communicative tasks. It is the speaker/writer who produces a text or an oral utterance of a certain form for the listener/reader to understand. Seen from this point of view, coherence is an aspect of the speaker/writers' *recipient design* that mirrors the fact that coherence is a factor of the comprehensibility of texts.[1] In the present book it is this perspective on coherence that I shall favour, without, however, losing sight of matters of understanding related to coherence or lack of coherence. Both coherence creating and coherence monitoring are relevant activities for successful communication.

---

[1] For the concept of recipient design, see Sacks (1992, vol. II, 230; 274).

The idea that the flow of human communication does not (always) proceed in a haphazard way but rather in (more or less) well-organised forms has been well-known since antiquity. Within linguistics, as I mentioned before, this idea became a popular topic of research in the 1970s, when linguists focused on the fact that some grammatical and lexical features not only contribute to the structure of the sentence but (also) to the connection between sentences. Favourite items of analysis for this strand of research were anaphoric pronouns and different types of "connectors", e.g., conjunctions and certain adverbs, which were considered to be the main means for making text "hang together" (e.g., Halliday/Hasan 1976). In the middle to late 1970s, however, various researchers, including myself, noted that these grammatical elements are neither necessary nor sufficient for creating the connectedness that is characteristic of coherent discourse (cf. Dascal/Katriel 1979, Hobbs 1979, Fritz 1982, 48-51). In addition, these authors – and others – indicated in which direction research on coherence should proceed. As Dascal and Katriel put it: "In their attempt […] to account for intuitions concerning the coherence and incoherence of discourse, researchers have generally recognized that they have to go beyond syntactic and semantic considerations, and incorporate a variety of pragmatic factors in their analysis" (Dascal/Katriel 1979, 76). Other linguists, of a more cognitive-functional persuasion, reflected on the relations between "the flow of thought and the flow of language" (Chafe 1979) and "coherence in text, [and] coherence in mind" (Givón 1993).

In some cases, scholars suggested that *one* single concept should play a fundamental role in the explanation of coherence, e.g., the concept of relevance (e.g., Wilson 1998), the concept of topic (e.g., Giora 1985, 1997), coreference (Grosz et al. 1995), or the assumption of "coherence relations" (e.g., Mann/Thompson 1988, Knott 1996, Bateman/Rondhuis 1997). Generally, however, studies in coherence have come to the conclusion that several kinds of resources are used in the achievement of coherence. This sometimes leads to the assumption of various "levels" of coherence (e.g., Moore/Pollack 1992, Stede 2007, 181ff.). The observation that seemingly heterogeneous aspects of language use are simultaneously effective in creating coherence presents multi-level approaches with a fundamental problem, namely the question of how to conceptualise the *interaction* of these heterogeneous factors and levels and their *integration* into a coherent theory.[2] This problem, again, leads to the general question of the kind of theory that one needs in

[2] Asher/Lascarides (2003), for instance, approach this problem by developing a modular theory of discourse interpretation.

order to model basic aspects of discourse coherence, including the complexity and adaptivity of forms of discourse. The present study aims at contributing to the solution of this problem by using a dynamic action-theoretic approach that integrates the study of basic resources for the construction of coherence and their implementation in discourse. In some places I shall refer critically to competing approaches (e.g., certain theories of "coherence relations"), but in general my aim is more constructive than polemical. And I am, of course, aware of the fact that there are still many puzzles of coherence to be solved.

Before going into more detail concerning the kind of theory I shall develop and explore in this book I want to draw attention to a very general matter of attitudes towards coherence. Simplifying the situation considerably, one could discern two basic attitudes to the study of coherence, the first aiming for the detection of (necessary) conditions that have to be met for texts and dialogues to be coherent, whereas the other is interested in the resources and strategies used for creating the various connections that contribute to coherence. While the former shows exclusionary tendencies and restrictive methods, concentrating on small well-behaved texts, the latter tries to be more realistic both with respect to a generous panorama of categories and the inclusion of "texts in the wild". My approach tends more to the second type, with occasional lapses into the opposite attitude. Attempting a realistic approach to the incredible variability and complexity of texts, it can be quite alarming to acknowledge the multiplicity of textual orders and the richness of combinatorial strategies, so a disposition to implement a restrictive regime in one's theory is understandable, if not necessarily successful. These rather general reflections will surface again at various points in the present book, for instance in those passages where I make a plea for a flexible theory of topic and a rich inventory of action categories.

As a preview to the topics of this book I will now list the main organising principles of discourse which I shall discuss and which can be considered the major *resources* for the achievement of coherence:

(i) speech act sequencing patterns (including commitment scores, utterance forms, coreference patterns etc.),
(ii) common ground and its dynamics,
(iii) topic structures.

In addition, I shall briefly discuss some of the principles speaker/writers and hearer/readers seem to follow when they create coherent discourse and try to make sense of a given discourse:

(iv) communication principles (e.g., principles of relevance, continuity, and consistency).

Changing the perspective to the *practices* involved in creating coherent discourse, the above resources are used in the following corresponding activities:

(i) the production of speech act sequences,
(ii) common ground management,
(iii) topic management.

On the part of the hearer/reader, the respective activities can be summarised as the *monitoring* of speech act sequences, common ground dynamics, and topic development.

What I shall try to demonstrate in several places is how the use of the different resources interacts in the production of discourse, e.g., how patterns of coreference interact with common ground dynamics and topic development or how speech act sequencing patterns are combined with topic structures to achieve longer passages of coherent text of a certain type.[3]

As for the characterisation of my theoretical approach as *action-theoretic* and *dynamic*, I should like to insert a few elucidating remarks at this point. The theory of linguistic action presupposed and further developed in this book is mainly inspired by Wittgenstein's concept of *language game*, which considers linguistic acts to be embedded in various types of practices, e.g., the practices of demanding an action and acting on the demand, describing an object, producing an object on the basis of a description, and telling a story (Wittgenstein 1958, § 23). Within such practices, the (linguistic) acts are serially organised, and it is this serial organisation that essentially contributes to the coherence of texts and verbal communications.[4] It is for this reason that what I call "sequencing patterns" play an important role in my account. In many cases, communicative practices are characterised by core speech acts, e.g., suggestions and objections in planning talks, assertions, objections, and arguments in controversies, and accusations and defensive moves in quarrels. Some types of complex (serial) acts can themselves be embedded in different types of practice, e.g., narratives, which can be part of

[3] The latter question has been discussed in some detail for various types of newspaper texts in Schröder (2003).

[4] A further concept in Wittgenstein's writings that is relevant to the problem of coherence is the concept of *Zusammenhang* (denoting connectedness, nexus, and context) (see Fritz 1982, chs. 1 and 4). For its usefulness in a theory of social practices, see Schatzki (1996, 14).

(historical) explanations, justifications of acts, and contexts of entertainment. And, finally, it is worth mentioning that in many practices different types of acts can be alternatively or additively performed, e.g., in teaching someone a certain gymnastics routine one can *describe* the routine, *physically demonstrate* it, and/or *show a film sequence* of somebody performing it.[5]

The last example also shows an additional advantage of an action-theoretic approach that is already indicated in Wittgenstein's examples of language games: linguistic acts are frequently combined with non-linguistic acts, e.g., physical acts like pointing at an object, demonstrating the use of an instrument or other acts involved in complex activities like cooking and teaching how to cook. In an action-theoretic framework it is quite natural to describe and analyse the interaction and coherence of these different kinds of acts and activities.

Concerning the consideration of aspects of discourse *dynamics*, there are two types of dynamics that play a role in this theoretical context. The first is the *micro-dynamics* of the flow of text or dialogue, which includes the serial performance of communicative acts, the incremental update of common ground, and the development of topics. As mentioned before, managing and monitoring these aspects of discourse dynamics are basic factors in the creation and understanding of coherence. It is therefore that these aspects will play a fundamental role in the following chapters. The second kind of dynamics concerns the historical evolution of communicative practices and individual aspects of these practices. Examples of such a kind of *macro-dynamics* are the historical evolution of the practice of scientific controversy since the 17th century, changes in the form of news reports since the first newspapers of the 17th century, and the evolution of narrative techniques in the 19th and 20th centuries. An important and difficult question in this context concerns the problem of how the routinisation of micro-dynamic strategies feeds into the macro-dynamics and how new and modified discourse practices emerge.[6] This second type of dynamics will only be touched upon in the present book.[7] One field where we can, even as lay persons, actually observe the relevant dynamics and at least some of the *micro-macro-links* is the development and change of public topics on the basis of the communication in

---

5 For the use of physical demonstrations, descriptions and pictures in teaching a sports routine, see Muckenhaupt (1986, 110-116).

6 It is worth noting that many of the early 17th century German news texts look fairly incoherent to modern readers. For an example, see Fritz (2017, 343f.).

7 For the history of the pragmatic organisation of controversies, see Fritz/Gloning/Glüer (2018). For the development of new text types in the early German newspapers, see Schröder (1995), Gloning (1996), Fritz (2001).

political debates, in the news media, in the social media, and in private conversations.

The idea that coherence is something produced by the speaker/writers raises the question in what way this production is directed or governed. A possible answer to this question could be that in producing coherent text speaker/writers follow a *principle* of coherence similar to Grice's Cooperative Principle and its accompanying maxims. Such principles have a normative aspect, which becomes apparent when we see people *criticizing* lack of coherence of a text or *complaining* of it. At the same time, it is not obvious that speaker/writers actually *know* this principle and its attendant forms of implementation in any detail. However, echoing Grice, we can say that following this kind of principle seems to be reasonable (rational), and speaker/writers routinely do so in order to achieve successful communication (see Grice 1989, 28-30). Accordingly, creating coherence seems to be, at least partially, a product of routines and of *knowing how*, which does not presuppose the ability to make explicit the principle and its forms of application. It is also worth noting that in general speaker/writers are *expected* to produce coherent contributions, so that a *presumption* of coherence is normally communicated by speaker/writers' utterances.[8]

Before approaching details of the kinds of resources and strategies speakers/writers use to produce coherent contributions to discourse I shall briefly focus on two more fundamental aspects of coherence:

(i) Coherence is basically a property of the *use* of *linear* sequences of linguistic material.
(ii) Coherence is a degree concept, viz., there are different degrees of coherence, from fairly strict coherence to loosely organised discourse.

As for the question of linearity, two comparisons might help to clarify this point. The first concerns a comparison with the concept of relevance. Whereas we can say that a statement made in a different text or at a distant point of a given text is *relevant* to the present argumentation in the text under scrutiny, we would not say that such a statement is textually *coherent* with the present argumentation. To be coherent (or incoherent), the two textual elements would have to be in a linear connection, for instance, in adjacent positions in the text. This should, however, not be taken too narrowly, as there are also discontinuous patterns of linear connection and the relevant connection maybe situated at the higher level of paragraphs and not of adjacent sentences. And it is indeed possible for elements of different texts to be se-

[8] See Sperber/Wilson's "presumption of relevance" (Sperber/Wilson 1986, 158).

quentially connected, e.g., in the case of controversies where a text T_2 may contain answers to accusations made in an earlier text T_1. In addition, it is worth mentioning that various non-sequential *forms of connectivity* can play a role in securing coherence, e.g., inferential relations, relations of consistency/inconsistency, the relevance relations mentioned above, certain forms of intertextuality, and multimodal connections (e.g., between text and image). These forms of connectivity will be treated in different chapters of this book.

A second comparison, which also introduces a special type of connectivity, concerns the *coherence of hypertext*. Hypertext is multilinear text, so, strictly speaking, the principle of coherence does not apply to hypertexts as such. It does, however, apply to individual *paths* through a hypertext (see chapter 7). Thus, hypertexts may be tightly organized and thereby show a high degree of connectivity, but they are not coherent in a primary sense.

As for the *gradation* of coherence, one can show that for some dialogue forms and text types, e.g., formal interviews, examinations, legal texts, and academic papers, coherence requirements are stricter than for others, e.g., informal conversations and narratives. This has to do with the functions of such discourses, which require a high degree of reliability, consistency and comprehensibility, among other qualities. It is, however, worth noting that (partial) lack of surface coherence is by no means an indicator of a low quality of a text. In many modern literary texts certain coherence principles are relaxed or intentionally flouted in order to create literary effects, as mentioned before.[9] In a way, such texts are inconsiderate, as they complicate normal forms of understanding, but "good literary texts are inconsiderate in an interesting way" (Zwaan 1996, 243). And one should not underestimate the joy of disorder. In Adorno's words: "The task of art today is to bring chaos into order." („Aufgabe von Kunst heute ist es, Chaos in die Ordnung zu bringen").[10] Thus, it is interesting in such cases to ask what the function of intentional incoherence (e.g., fragmentation) is and what role the reader's looking for coherence plays. Even in academic writing there are texts that do not conform to traditional principles of connectivity and will still be considered texts of the highest quality, e.g., Wittgenstein's "Philosophical Investigations" with their sometimes seemingly unordered sequences of "remarks". Therefore, a high degree of coherence may be a criterion of quality for some texts and not for others.

---

[9] See, for instance, Peter Handke's reflections on the "unconnected coexistence ('unverbundenes Miteiander') of many different forms of writing" in one text, with fragmentary elements and seamless transitions (Handke 1990, 68f.), and his own practice of writing in his "Versuch über die Jukebox".

[10] Adorno, "Minima Moralia", aphorism(s) No. 143.

It is interesting to see how authors sometimes struggle with problems of coherence and, for various reasons, settle for less coherent texts than could have been imagined. An example can be given from Shwayder's "Stratification of behaviour", where the author at one point feels he ought to give some hints towards the solution of a problem for which he cannot provide a systematic account (in this context): "But some elucidation is in order here, and I should like to make a few rather disconnected observations here for that purpose" (Shwayder 1965, 41). "Disconnected observations", so the author assumes, can at least achieve some clarification as a first step towards a satisfactory treatment of a "challenging issue". A different kind of reason for choosing a somewhat "disconnected" form of representation is given in Roland Barthes' "Notes on André Gide and his *Journal*" (Barthes 1942/1985). As a preface to his notes Barthes presented the following reflections:

> Held back by the fear of enclosing Gide in a system of which I knew I could never be satisfied, I searched in vain for a connecting link for these notes. On reflection, it is better to present them as they are, without trying to mask their discontinuity. *Incoherence seems to me to be preferable to an order that deforms* (Barthes 1942/1985, 85; my translation; italics added).
>
> [Retenu par la crainte d'enclore Gide dans un système dont je savais ne pouvoir être jamais satisfait, je cherchais en vain quel lien donner à ces notes. Réflexion faite, il vaut mieux les donner telles quelles, et ne pas chercher à masquer leur discontinu. L'incohérence me paraît préférable à l'ordre qui déforme (Barthes 1942/1985, 85).]

For his particular purpose in writing these "notes", Barthes obviously preferred an "open" list structure to imposing a systematic treatment on Gide's journal, which might "deform" its object. Still, in reading the notes, which are internally quite coherent, one finds various connections between the fragments, so that a kind of (partial) order emerges after all.[11]

These reflections on forms of coherence and incoherence indicate that the concept of coherence should be regarded as a prototype concept, i.e., it appears difficult to formulate necessary and sufficient conditions for coherence, whereas we can easily show typical examples of coherent discourse, less typical examples, and marginal examples. As opposed to the rules governing a game like chess, the "rules" of coherent discourse are flexible, strongly contextualised, and can sometimes be formulated in terms of preferences. Attempts at formulating strict "constraints" on coherence are very instructive

[11] An arrangement of "fragments" as a form of representation was used by Barthes also in other works, e.g., his "Fragments d'un discours amoureux" (Barthes 1977).

in this respect, a point I shall discuss in section 8.4 of this book.

Utterances which might look incoherent on the surface but which are of quite a different type and origin as the ones discussed so far are to be found in the discourse of small children and patients with aphasic disorders or schizophrenia. Whereas small children (up to about age three) have not yet acquired certain phonological, lexical and grammatical means and certain routines of text production, aphasics, depending on the type of brain lesion responsible for their impairment, have (partially) lost the use of some phonological, grammatical or lexical means, which reduces their ability to create superficially coherent text. This does, however, not necessarily mean that they cannot make themselves understood and successfully participate in communication. In many cases, if there is enough common ground with their interlocutors and if these are prepared to use their interpretive ability, such "unorthodox" utterances can be understood as coherent contributions to the ongoing communication. Especially in the case of aphasics, such utterances can sometimes be interpreted as *attempts* to use more standard expressions. What these observations show is that the speaker's intention to express something and the recipient's readiness to decipher this intention may take priority over the presence of superficial features of cohesion.[12]

On the other hand, the assumption of incoherence in a given discourse can be a strong indicator of lack of understanding, lack of common ground, differences in the perception of topics, and, generally, of a breakdown in communication. Thus, a *prima facie* diagnosis of incoherence may lead to doubt as to the chances of successful communication, but it may also prompt more persistent efforts of interpretation.

Another aspect of coherence that can lead to different forms and degrees of coherence is the difference between *local* and *global* coherence. Local coherence relates mainly to sequences of utterances used in restricted contexts, e.g., within adjacent utterances in a dialogue or within a paragraph of a written text, whereas global coherence concerns the connections between such smaller units in larger stretches of discourse, which may, for instance, be closely linked thematically or less so. Thus, it is possible for a text to be locally coherent and globally non-coherent or less coherent and vice versa. In addition, a discourse may be more strictly organised in some passages and less so in others.

---

[12] For reflections on coherence judgments and the interpretation of aphasics' and children's discourse, see Pulvermüller (1990, 78f.), Linnik et al. (2022); Keenan/Klein (1975), and Biere (1978, 192-195). On the problem of "disordered discourse" in schizophrenia, see Gernsbacher/Tallent/Bolliger (1999).

Setting aside these special cases and returning to the beginning of this introduction, the general question we have to explore seems to be the following: What are the resources that, taken together, make it possible for speaker/writers to perform (for their audiences) their communicative task of creating coherent dialogue and text adjusted to widely varying purposes, often responding robustly to surprises, and to produce the astonishing complexity of the world of discourse.

I now turn to a question about the structure of the field of discourse studies that plays an important role for the textual organisation of the present book. This question concerns the differences and similarities of the resources and practices of coherence construction in (mainly spoken) dialogue and (mostly written) monological text. There seems to be at present a strong current of thought favouring the assumption that these resources and practices are fundamentally different. However, not all linguists agree with this stance. Givón, for example, recently characterised the situation as follows:

> The tradition of Conversational Analysis (CA), imported into linguistics from sociology, holds that the coherence structure of conversation, where the control of perspective shifts periodically from one speaker to another, is radically different from the coherence structure of narrative or procedural discourse. Several studies have suggested that this claim is not empirically supported (...)" (Givón 2020, 106).

I think that Givón is right on this point, and I myself would like to show that fundamental resources of communication are used in both dialogue (aka conversation) and monological text, but that in some, albeit not all, respects the practices and strategies of *using* these resources are different and should be analysed accordingly.[13] It is, of course, true that some practices like turn-taking, using joint attention as a resource of common ground, the use of prosodic means, certain so-called "repair" formats, and certain forms of the emergence of topics are typical of spoken dialogue and form important issues of analysis. On the other hand, many "monological" sequences like "claim & arguments for the claim" or longer passages of narrative occur both in spoken dialogue and monological text, common ground is used as a resource in both kinds of communication, and certain topic structures are used for topic management in both. And, of course, there are many formal types of dialogue, e.g., interviews, examinations, court hearings etc., where longer passages of "text" are spoken by one speaker without turn-taking. Inversely, many written texts are produced in the context of dialogical interaction, e.g.,

[13] See also Asher/Lascarides (2003, 293ff.): "Why dialogue and monologue are similar".

in the case of email-exchanges or in the course of controversies conducted in writing, where the coherence structures of the relevant texts are fundamentally determined by the position of the respective text in the controversy. It is partly on account of these observations that I tend to view monological text (also) dialogically and dialogues (also) textually.

On a more general level, it seems to make sense from an evolutionary point of view to assume that resources used for one form of communication should also be available in the other, as, in many cases, the communicative tasks to be solved are similar.

Independently of the fact that the view expressed just now may have a certain plausibility, which I hope it has, it presents quite a problem for the production of a unified text. Having to go back and forth between specific problems of dialogue and of text makes for a fairly awkward flow of text, whereas by producing two clearly separated "blocks" on text and dialogue one risks missing important commonalities. An extreme solution to the problem would consist in writing two books, one on "text" and one on "dialogue", with extensive cross-references. This, in fact, was the solution chosen by the editors of a major handbook on discourse analysis (Brinker et al 2000/2001), who decided to produce two volumes, one on "text" and one on "conversation" ("Gespräch"). As the opportunity to give extensive cross-references between relevant chapters of the two volumes was missed, the handbook potentially contributed to obscuring the common features and to reinforcing a problematic demarcation between two fields of research.

For the present book, I decided to treat both topics together, starting with a number of chapters dealing mainly with "text", but introducing in general terms the main resources used to create coherence and giving cross-references to special aspects of dialogue where it appeared necessary or useful. In the second part of the book, I introduce basic features of dialogue and aspects of the creation of coherence in dialogue, making use of the categories introduced earlier. To some readers this order of topics may appear strange, as there are good reasons for assuming that dialogue is in a sense more fundamental than written text production. To this objection I have two answers. In the first place, I found it easier to introduce the basic aspects of coherence without having to tackle the special problems of turn-taking etc. And secondly, as I mentioned before, the kind of theory I present is in itself "dialogical" in the sense that it assumes fundamentally dialogical resources and practices like the use of common ground, the commitments of speaker/writers, and the practice of audience orientation, both in text production and dialogue. Thus, the seeming gap between treatments of text and dialogue is much reduced and both kinds of the use of language can be viewed from a common perspective.

2. Basic discourse units

2.1 Communicative tasks and functional building blocks

When writing a text or participating in a conversation, we have to fulfil various communicative tasks, depending on the type of text or conversation. Thus, communicative tasks are frameworks for coherence. We may have to address the audience, identify and describe certain objects, report certain events, evaluate actions, and summarise our claims – just to name a few. We fulfil such tasks mainly by using linguistic expressions of varying size and kind. In some cases, we also use pictures, graphs or sound bites. The textual and conversational elements we use to fulfil these tasks we shall call basic discourse units or, if we look at their use, functional building blocks.

The following are a number of basic discourse units which regularly play an important role as functional building blocks in the organisation of a discourse:

(i) subsentential units (phrases, constructions etc.)
(ii) sentences/clauses
(iii) sentence sequences
(iv) paragraphs and dialogue sections
(v) text/image configurations

In the following sections I shall discuss these units and show some of their basic functions.

In addition, I shall briefly introduce the concept of discourse type. Discourses of certain types serve to fulfil more comprehensive communicative tasks, and a given type provides a general outline for the global organisation of functional building blocks. Hence, I add here as an extra unit:

(vi) discourse types (text types and dialogue types).

2.2 Subsentential units and their functions

Among subsentential units, phrases often have a distinctive communicative function. Arguably the most important type of phrase for the coherence of discourse is the noun phrase (NP), including proper names, definite and indefinite descriptions, and personal pronouns. In the first place, noun phrases

are used to identify and re-identify objects (1); other functions include naming a topic (2), for instance when announcing it (3), as in the following examples:

(1) [The house]$_{NP}$ stood on a slight rise just on the edge of the village. [It]$_{NP}$ stood on its own and looked out over a broad spread of West Country farmland. Not a remarkable house by any means – [it]$_{NP}$ was about thirty years old, squattish, squarish, made of brick, and had four windows set in the front of a size and proportion which more or less exactly failed to please the eye. (Douglas Adams, *The Hitch-Hikers Guide*, 9)

(2) On [the requirement that transformations preserve meaning]$_{NP}$ (The topic of a paper: Barbara Hall Partee 1971)

(3) I shall now continue by talking about [counterfactual conditionals]$_{NP}$

What makes noun phrases so useful is their potential for structural complexity, which allows the speaker/writer to make subtle differentiations of objects to be identified and fulfil extra tasks like, for instance, indicating common knowledge, as in (4):

(4) [The first of the complicated formulas we discussed last week]$_{NP}$

In this case the speaker/writer identifies one out of a lot of formulas by indicating its place in a list and its property of being complicated and by reminding the audience that it was discussed a week before.

A second type of phrase worth mentioning is the prepositional phrase (PP). PPs are often used as adverbials to indicate time frames and places, like in the following two examples from Steven Nadler's biography of Spinoza: [1]

(5) (a) [Throughout 1676]$_{PP}$, Spinoza carried on his philosophical exchanges with Tschirnhaus and Oldenburg and continued to work on notes for an amended edition of the *Theological-Political Treatise*. […] (b) He probably did not travel very much during this year, (c) although he continued to take time out from this work to entertain numerous visitors. (d) Among the guests who dropped by the house on the Paviljoensgracht was Leibniz, the German philosopher-diplomat who had written to him five years earlier […]

With the adverbial PP *throughout 1676* in the initial position of (5a) the author provides a time frame that is active for a longer stretch of text. This time frame helps us to understand what *during this year* in (5b) means and it also permits us to understand sentence (5d) in such a way that we now know that Leibniz's visit occurred in the same year, although this is not explicitly mentioned. And we can also figure out that Leibniz wrote to Spinoza in 1671

[1] Nadler (1999, 340 and 212).

(*five years earlier*). Earlier in the book we find a passage about the plague, which returned to Europe in 1663. The part of the story that is situated in Amsterdam is opened up by (6), where *in Amsterdam* has the function of providing a local frame for this part of the narrative:

(6) [In Amsterdam]$_{PP}$, where the plague reportedly began, there were almost ten thousand deaths in 1663.

This technique of creating local and temporal frames for the following sections of a story is quite frequent in narratives where a certain series of narrated events continues to happen during the same time span and at the same place, allowing the author to assume this continuity without having to repeat local and temporal reference explicitly.

When analysing texts like, for instance, news reports one also frequently finds typical "chunks" of expressions combining phrases of these types, e.g., expressions like *the fight against* NP, where the position of NP can be filled by various nouns, depending on the topic of the report:[2]

(7) the fight against terrorism/apartheid/climate change/corona/cancer/depression/corruption

Obviously, this kind of expression is particularly suitable to express a dramatic, warlike situation in such reports. The same is true of semantically related expressions like *the war against* (*terrorism*, *drugs*, *nature*) or *the war on* (*drugs*).

Recurrent sequences of three or more words like *I don't know what*, sometimes called "lexical bundles" (Biber et al. 19995, ch. 13.2), are particularly frequent in conversational discourse, serving as typical building blocks.

Another type of subsentential unit frequently used in dialogue are contextually licensed short syntactic forms, like in the following example:

(8) A: What are you looking for?
B: My glasses.

I shall return to this point in the following section.[3]

Generally speaking, all types of subsentential expressions can play a functional role in discourse by contributing to the meaning of a total utterance: Adjectives like *beautiful* or *absurd* can be used in evaluative utterances, adverbs like *probably* and modal verbs like *must* and *may* in their "epistemic" uses can indicate the speaker's degree of certainty regarding the truth of a

---

[2] A Google search provides 106 million instances of the unit *fight against,* which shows its currency.

[3] Some more details of the language of dialogue are given in section 12.4.

statement, discourse markers like *well* can be used to indicate doubt or hesitation (among other functions), *that*-clauses like *that it is raining* can be used to express propositional content, and so on.[4]

Finally, among subsentential units that play a role for the coherence of texts, I mention here constructions which trigger knowledge frames for the reader to anticipate further developments of the text. Typical expressions of this type are verbs like *to buy* which indicate a kind of *scene* or *frame*, including a seller, a buyer, and a merchandise being sold:

(9) (a) Peter bought a new bicycle last week. (b) The shop assistant told him he could return it for servicing any time.

The buying frame indicated by the use of *bought* in (9a) can be used to introduce the shop assistant in the following sentence (9b) and in this way to connect the two sentences. This is a case of what has been called "bridging" (see Clark 1975, Asher/Lascarides 1998).

2.3 Sentences and communicative acts

2.3.1 Communicative acts

Sentences have traditionally been considered the basic units of texts. In the framework of an action-theoretic approach this idea has to be modified. In such a framework it is *communicative acts* that play the role of a fundamental text unit, and communicative acts are standardly performed by the *use* of sentences/clauses (and their counterparts in dialogue). Of course, communicative acts can also be performed by using pictures, as I shall discuss later on. In this section, however, I shall concentrate on *linguistic acts* or *speech acts* and the sentences and other linguistic forms used to perform them.[5]

Taking communicative acts as a basic unit of discourse has the further advantage that it is quite natural to include certain types of subsentential utterance forms that can be used to perform speech acts, e.g., certain dependent clauses (*because it is raining*) and clause-like PPs (*because of the rain*) as well as so-called supplementary relative clauses like the following: *I gave it to John*, *who passed it on to Mary* (see Huddleston/Pullum 2002, 1064). The

[4] Subsentential units that are used to perform complete speech acts will be mentioned in 2.3.1.

[5] Alston (1964b) uses the expression *linguistic acts*, Austin and Searle prefer *speech acts* or *illocutionary acts*.

same is true of the many types of subsentential forms used to perform speech acts in dialogue, e.g., noun phrases in answers to *who*-questions (*my brother*), truncated sentences like *I didn't* as reactions to accusations, and exclamations like *Fucking stupid*.[6]

The kind of theory of linguistic acts that I presuppose in the following discussion diverges quite substantially from traditional speech act theory as propounded by Searle and his followers (e.g., Searle 1969). Three differences are worth noting from the beginning:

(i) The structure of speech acts is more complex than assumed in Searle's theory.
(ii) There are act types that differ from typical illocutionary act types as described by Searle, e.g., act types like "preparing a narrative by describing a person", "repeating a request", "insisting on a claim", "rebutting an objection".[7]
(iii) Speech acts are not considered as isolated acts but as parts of speech act sequences.

2.3.2 Speech acts

A description of a speech act consists mainly of the following elements:

(i) a speech act characterisation (e.g., "speaker A asserts"),
(ii) a specification of the propositional content (e.g., "that it is raining"),
(iii) a specification of the form of utterance used in performing the speech act (e.g., "A asserts that it is raining *by* uttering *it is raining*"),
(iv) a specification of a *by*-relation (when applicable) (e.g., "A accuses B of x-ing *by* asserting that B has x-ed")
(v) a specification of the commitments the speaker undertakes by performing the speech act,
(vi) a specification of the linear context in the text flow in which the act is performed (e.g., relevant acts performed before and after the act described).
(vii) (when applicable) a description of a communicative act that is performed simultaneously with the given speech act (e.g. by simultaneously showing an illustration)

---

[6] This exclamation is taken from a conversation I discuss in chapter 13.
[7] For lists of "classical" speech act types and so-called performative verbs, see Searle (1979, ch. 1) and Vanderveken (1990, ch. 6).

At this point a few explanations are in order. Of these parts of the description of a linguistic act, the first three are standard elements of the speech act rules described in speech act theory. Therefore, I shall say no more about them.

Element (iv) is intended to model the fact that in many cases a speech act like assertion is used to perform another act, e.g., to accuse somebody of something, like in the example given above, or to evaluate an object: By asserting that the ring is made of gold the speaker characterises it as valuable. (iv) was originally not used in speech act theory, apart from mentions in passing (e.g. Austin 1962, 121ff.), but it was frequently discussed in the (general) theory of human action.[8] A typical example of what Goldman calls "level generation" is "*by* extending his arm out the car window S is signaling for a turn" (cf. Goldman 1970, 25).[9] The *by*-relation was explicitly introduced into the description of speech acts in Heringer (1978) and is used by me for the many cases where in discourse an act of x-ing is performed by way of another act, mostly an assertion. (In descriptions of such acts with *by*-relations I shall write the *by* in italics.) Further examples are: Speaker A can *answer* the question, where they have been, *by* stating that they were in the garden.[10] Or: A speaker/writer can *explain* a request for help *by* stating a reason why they cannot fulfil a given task themselves:

(10) Could you look after the children tomorrow morning? I have to see my doctor.

For the analysis of text coherence, one of the useful features of this form of speech act description consists in the fact that it explains why in texts the use of declarative sentences frequently plays an additional role to the role of assertion they are standardly assumed to perform.

2.3.3 Speech acts and commitments

Element (v), the specification of relevant *commitments*, is an alternative to the specification of so-called "preparatory conditions" or the "sincerity condition" in Searle's speech act theory (Searle 1969, ch. 3). This idea was already present in an early version of speech act theory presented by Alston:

---

8 E.g., Anscombe (1957, 23ff.), Shwayder (1965, 134), Goldman (1970, ch. 2).

9 In fact, Goldman's relation of "level generation", symbolized as "→", is the inverse relation of the *by*-relation. But I shall not discuss this difference here. For a brief history of the use and discussion of the *by*-relation, see Fritz (2017, ch. 2.3).

10 Here and in most of my text I use the "singular *they*".

"What is required for a given illocutionary act, in addition to the utterance of an appropriate sentence, is not that certain environmental conditions actually hold or even that the speaker believe them to hold, but only that he take responsibility for their holding" (Alston 1964a, 42f.).

Alston's concept of responsibility has a *dialogical* aspect that is important for a theory of human communication: „Responsibility for x being the case is essentially connected with the possibility of being called to account if x is not the case, and such a possibility can be taken as an indication of responsibility" (Alston 1964a, 41)

To give an example: If speaker A makes the assertion that p, he undertakes the *commitments* that he believes that p and that he is prepared to present evidence for p on demand.[11] In the case of a *by*-relation additional commitments are incurred by the speaker/writer as the following description of the action pattern of "complaining" shows:[12]

> A can complain that p *by* asserting that p *by* uttering *p*.
> In doing so A undertakes the commitments
>
> (i) that A believes that p,
> (ii) that A is prepared to present evidence for p,
> (iii) that A is discontent with p,
> (iv) that A has reason to be discontent with p.

In this case, (iii) and (iv) are the additional commitments characteristic of the complaint made *by* the assertion that p.

For all these commitments, which, unless explicitly cancelled, are valid throughout the following discourse, the audience can ask A to make good a commitment by, for instance, presenting evidence for p or explaining why they are discontent with p.

Commitments are both "forward-looking" and "backward-looking".[13] They are forward-looking in the sense that they play a role in determining what is acceptable and non-acceptable in the further course of the text or dialogue. If, for example, a writer complains about p, thereby committing

---

[11] For a brief history of the concept of commitment, see Fritz (2017, ch. 2.4). For a recent discussion of the role of commitments, see Beurts (2019); Krifka (2019).

[12] Where I want to emphasise that I am speaking of a speech act *pattern*, I shall put the descriptive expression in double quotes, like in "complaining" above.

[13] For the special case of assertions/statements Dummett formulated a similar feature as follows: "Learning to use a statement of a given form involves, then, learning two things: the conditions under which one is justified in making the statement; and what constitutes acceptance of it, i.e., the consequences of accepting it" (Dummett 1973, 453). For the role of "discursive commitment" and "entitlement", see also Brandom (1994, 159ff.).

themselves to their discontent with p, they can go on to explain *why* they are discontent with p. So the commitment undertaken with the complaint opens up a specific option for the coherent continuance of the text. But it also precludes certain further acts performed by sentences of the text, e.g., an assertion of satisfaction with p. Commitments are backward-looking (like presuppositions) in the sense that a complaint about p cannot generally be coherently uttered if in an earlier part of the text the writer asserted their enthusiasm for p. (In an improvised speech, for example, it may however be possible for the speaker to change their mind and *retract* their expression of enthusiasm and then utter his complaint. But this would be an unusual and therefore noteworthy occurrence.) Thus, monitoring the consistency of the acts and their commitments is part of the writer's coherence management. And the same applies to dialogues, about which I shall say more later.

An interesting effect of the undertaking of commitments is the fact that if an aspect of an assertion that had not previously been part of the common knowledge of the participants – e.g., writer and reader – is introduced into the discourse in the form of a commitment, this aspect automatically becomes part of the common knowledge of the participants. For example, if A states that his sister has recently given up smoking and his addressee B up to this point neither knew that A had a sister nor – a fortiori – that she had smoked until recently, then these two facts automatically become part of the set of commitments of A and of the common knowledge of A and B. This effect, which is explained quite naturally by the concept of commitment, has been much discussed under the heading of "accommodation of presuppositions" (see Lewis 1979).

This observation also leads us to another aspect of the concept of commitment. Following Hamblin's "formal dialectic" (1970, 263ff.), we can assume that after performing a communicative act the relevant commitments are added to the speaker's *commitment store*. In the course of a sequence of utterances in a discourse the speaker/writer's commitments accumulate so that the dynamics of the commitment store represents a (partial) history of the acts performed with the sentences of our text. One could also model this aspect of the dynamics of a discourse in terms of a game played by the speaker/writer in making their utterances and term this documentation of commitments the "text score" or, in the case of dialogue the "conversational score" (cf. Lewis 1979, 176). Details of the commitment store and its role in the dynamics of common ground will be discussed in the following chapter.

2.4 Sentence sequences and speech-act sequences

The idea of the "forward-looking" and "backward-looking" character of commitments already presents a first kind of connection between linguistic acts in a sequence of acts that contributes to the creation of coherence in the sequence.

There is, however, another type of connection I should like to mention at this point, which I shall go into in more detail in the chapter on "sequencing patterns" (ch. 4). This aspect of speech acts, which fundamentally contributes to the coherence of discourse, was mentioned as element (vi) in the list of aspects of the description of linguistic acts above. Some linguistic acts, performed by adjacent sentences, show a particularly close functional connection to their neighbours in the text sequence or dialogue. Some acts of this type can be generally characterised as acts *supporting* their preceding or following act. Sometimes this support takes the form of an *explanation*, in other cases, the point of the act consists in *preparing* a following act, for example in the case where the speaker/writer gives some information which is necessary or at least useful for understanding the following act. If, for instance, A wants to make a remark about the importance of definite descriptions, a term that might not be known to their addressee, they can prepare their remark by giving an example of this type of expression like in the following sequence:

(11) (a) Expressions like *the present king of France* are called *definite descriptions*. (b) Bertrand Russell considered definite descriptions so important, philosophically, that he wrote a chapter about them when in prison for his pacifist views in 1918.

The same information can also be given as a supplementary appositive, as in (12), or as an addition, following the "main" act, as in (13), although the latter could be considered a bit awkward:

(12) (a) Bertrand Russell considered definite descriptions, i.e., expressions like *the present King of France*, so important, philosophically, that he wrote a chapter about them when in prison for his pacifist views in 1918.

(13) (a) Bertrand Russell considered definite descriptions so important, philosophically, that he wrote a chapter about them when in prison for his pacifist views in 1918. (b) Definite descriptions are expressions like *the present King of France*.

Another frequent type of supporting move *following* the main act is an explanation of the motivation for the preceding act, like in the example of a request plus an explanation I gave earlier on in connection with the introduction of the concept of *by*-relation:

(14) (a) Could you look after the children tomorrow morning? (b) I have to see my doctor.

For some of these sequencing types there are connectives which serve to make explicit the kind of connection between the two acts. Example (15), for instance, is similar to example (14), but in this case the explanation of the situation stated in (15a) is added using the connective *because* (15b):

(15) (a) I cannot look after the children tomorrow morning, (b) *because* I have to see my doctor.

Generally speaking, supporting acts preparing or explaining another act are frequent not only with individual simple acts like in examples (10) to (15), but also in connection with longer spans of text like narratives or teaching materials. In the case of narratives, we frequently find acts of "orientation" at the beginning of the narrative, where information is given about persons, places and historical periods that play a role in the narrative. Of course, this kind of information can also be given piecemeal in the course of the narrative "just in time" when it is needed. In the case of teaching materials, authors sometimes provide what is called an "advance organizer", i.e., "background material of a conceptual nature presented at a much higher level of generality, abstraction, and inclusiveness" (Ausubel 1960, 271) that is meant to improve access to the material it precedes by explaining and organising the basic ideas of the material it precedes.

Whereas simple explanations of the type shown in (10) to (15) are closely related to traditional types of illocutionary acts, the last examples show that there are also more complex types of acts used for the preparation and "post-processing" of other spans of text, e.g., complex explanations, descriptions, evaluations, and summaries that are used to provide a more detailed or more specific understanding or a condensed version of the text passages they refer to. I will talk about these "big" speech acts in section 4.10.2.

Finally, I now briefly address a last aspect of the description of linguistic acts – (vii) in my list of aspects –, namely the case where an additional communicative act is performed simultaneously with a given speech act. In the case of an illustration one can make a statement about an object *and simultaneously* present a picture of this object or an object of this kind. In some languages there are also linguistic means for the simultaneous performance of two (or more) speech acts, e.g., in the case of an evaluation of a person made with a non-restrictive relative clause in the course of an assertion:

(16) John, who is otherwise quite nice, tends to forget these things.

As some of the act patterns discussed in the last paragraphs, e.g., asserting and explaining the assertion, are routinely connected to others in a sequence

of acts, I shall come back to some of them in chapter 4 on "sequencing patterns".

2.5 Paragraphs and dialogue sections

Paragraphs are middle-sized textual building blocks which form functional and/or thematic units. Units of this type not only play an important role in traditional linear texts but also in hypertexts which can be described as clusters of paragraph-sized elements connected by links. It is *within* paragraphs that local coherence is most strictly organised. In well-organised written texts, paragraphs are frequently indicated by headlines, indention or other cues. Paragraphs normally begin at points in a text like the following:

1. At the beginning of *a new functional unit*
1.1 At the transition from a description of an instrument (in a technical manual) to a tutorial on how to use the instrument
1.2 At the transition from the presentation of a theoretical position to the arguments for this position.
1.3 At the transition from the narration of an event to the explanation of this event.

2. At the beginning of *a new thematic unit*
2.1 At the transition to a new subtopic
2.2 At the transition from a higher-level topic to a lower-level topic
2.3 At the transition from one event to the next in a narrative

This shows that paragraphs are important units for functional structuring and topic management.

It is worth noting at this point that paragraph-like structures also occur in dialogue. They are bounded units that serve to deal with a joint communicative sub-task by performing a series of speech acts focused on a certain (sub)topic. In formal dialogue types like examinations such units can be quite noticeably structured, whereas in informal conversation the flow of speech acts and topics is generally much more loosely organised, with subtle shifts leading from one (possible) section to the next. Partly due to this more fluid organisation, there is, as opposed to the concept of paragraph, no generally accepted term for such middle-sized sections of joint communicative activities. They have sometimes been called *phases* or *episodes* (Linell 1998, 183), but I shall just speak of *sections* where it seems necessary.

The various techniques of linking one paragraph to the next are among the most salient means for combining local and global coherence. *Implicit linking* relies on the common knowledge of functional and thematic relations like the ones just mentioned. For instance, knowing that topic T_2 is a subtopic of

T_1 "licenses" the transition from paragraph 1 to paragraph 2. Similarly, the knowledge of the relation between narration and explanation of an event licenses the transition from a narrative paragraph to an explanatory one. On the other hand, we frequently find forms of *explicit linking*, e.g., beginning a new paragraph with an explicit cross-reference to an aspect of the preceding paragraph (e.g., *This tension between …*) or by using a connective (e.g., *Texts can therefore …* or *Thus*, *the genre of …*).,

The following example shows how paragraphs are used for topic management in texts,[14] e.g., for introducing and establishing a topic and introducing a subtopic. It is taken from the so-called "lead section" of a Wikipedia article, which is intended to serve "as an introduction to the article and a summary of its most important contents".[15]

(17) (a) **Glenn Herbert Gould** (/gu:ld/; 25 September 1932 – 4 October 1982) was a Canadian classical pianist. (b) He was one of the best known and most celebrated pianists of the 20th century, and was renowned as an interpreter of the keyboard works of Johann Sebastian Bach. (c) His playing was distinguished by a remarkable technical proficiency and a capacity to articulate the contrapuntal texture of Bach's music.

(d) Gould rejected most of the standard Romantic piano literature by Chopin, Liszt, Rachmaninoff, and others, in favour of Bach and Beethoven mainly, along with some late-Romantic and modernist composers. (e) Although his recordings were dominated by Bach and Beethoven, Gould's repertoire was diverse, including works by Mozart, Haydn, and Brahms; pre-Baroque composers such as Jan Pieterszoon Sweelinck, William Byrd, and Orlando Gibbons; and 20th-century composers including Paul Hindemith, Arnold Schoenberg, and Richard Strauss.

The general topic of the article from which these paragraphs are taken is the person Glenn Gould, and the first paragraph introduces Gould as a famous classical pianist and establishes this topic by giving basic information on his playing. The first sentence (a) begins by giving his name, the pronunciation of his name, his life dates, and his profession. The sentences (b) and (c), each referring to Gould by a pronoun (*he* and *his*), continue to establish the topic by giving information on what he was famous for. The option of referring again – maybe several times – to a person introduced at the beginning of the paragraph by using pronouns instead of a proper name (or a definite description) is frequently used *within* paragraphs, whereas in starting a new para-

[14] Topic management in text and dialogue will be dealt with in more detail in chapters 5 and 14.

[15] https://en.wikipedia.org/wiki/Glenn_Gould (with small alterations, e.g., replacing the links by "normal" words). For the function of the lead section, see https://en.wikipedia.org/wiki/Wikipedia:Manual_of_Style/Lead_section.

graph writers tend to prefer to use a definite description or the proper name – as in our example.[16] The second paragraph – sentences (d) and (e) – introduces a subtopic of the previous topic, namely Gould's repertoire, and describes this repertoire by naming relevant composers and periods (Romantic, pre-Baroque).

If we look at the general structure of the Wikipedia article we find a number of major sections – a kind of big paragraph – with individual headings indicating their topics, e.g., his life, his recordings, and the awards he received, which are again divided into shorter paragraphs dealing with his early life, his eccentricities and other relevant topics. Thus, the article, as is usual with Wikipedia articles, has a hierarchical topic structure with paragraphs of different sizes as its basic structure and, in addition, links to subtopics of the article's topic, which create a network structure.[17]

2.6 Text/image configurations

Many texts we encounter today use images as functional building blocks. This is true, for instance, of advertisements, newspapers, textbooks of medicine, technical manuals, assembly instructions, scientific papers, Powerpoint presentations, and travel brochures. Authors producing such documents carry out communicative tasks by using different means (or "modes") of communication, namely text and image(s), in various constellations. Their actions performed in using these means can be described as a combination of *telling* (by verbal means) and *showing* (by pictorial means).[18] This practice serves to expand their scope of action by affording them options for the performance of communicative acts which cannot be performed by either text or image alone. Because this kind of communication uses different "modes" it has been called "multimodal". Multimodal communication also includes the use

[16] This option will be dealt with in more detail in section 4.9 on reference and co-reference.

[17] For general discussions of the status of paragraphs, see Longacre (1979), Givón (1983, 9-15), Stark (1988); Hearst (1997). For the analysis of Wikipedia articles in the framework of discourse analysis, see Gredel (2017).

[18] The following account is based on the action-theoretic approach to the analysis of text and image developed in Muckenhaupt (1986). In addition, I draw on work by Bucher (e.g., Bucher 2017), who analysed the reception of multimodal documents from an action-theoretic point of view.

of resources like colour, typography, page design, diagrams, sounds, video sequences, hypertext etc.[19]

In order to analyse how text and image can be combined to create coherent text/image documents it is helpful to start by looking at the way images alone can be used. Generally speaking, there is a fundamental similarity between the uses of sentences/texts and pictures, which consists in the fact that both can be used to carry out communicative tasks. To illustrate this, I now draw on a thought experiment performed by Wittgenstein in his "Philosophical Investigations":[20]

> Imagine a picture representing a boxer in a particular stance. Now, this picture can be used to tell someone how he should stand, should hold himself; or how he should not hold himself; or how a particular man did stand in such-and-such a place; and so on. [...]

For vividness and to aid the imagination, I insert here the following picture as an illustration:[21]

Described as speech act types, the uses Wittgenstein mentions here are the following:

(i) A instructs B how he should hold himself *by* showing the boxer picture.

---

[19] On the use of these different resources in multimodal city and country portraits, see Ermakova (2015). For an introduction to the analysis of multimodal discourse from the point of view of semiotics, see Kress/Van Leeuwen (2001). For a discussion of problems and methods in the analysis of multimodality, see (2014) and Bateman et al. (2017).

[20] This remark (probably) belongs to § 22 of the "Philosophical Investigations", where Wittgenstein reflects on Freges concept of "Annahme" ("assumption").

[21] In its original context, this photograph was used as an attention getter on a website dedicated to boxing training (https://boxing-arts.com). (21.04.21).

(ii) A instructs B how he should *not* hold himself *by* showing the boxer picture.
(iii) A reports to B how a particular man stood in such-and-such a place *by* showing the boxer picture.

What we find in this short list are two different illocutionary types, namely "instructing" and "reporting", and three different propositional contents, namely "how he should/should not hold himself" and "how the man stood". One could easily extend this list by the following items:

(iv) A identifies a particular boxer *by* showing the boxer picture.
(v) A informs B about what A's brother looks like *by* showing the boxer picture.
(vi) A introduces the topic of boxing *by* showing the boxer picture.

In order for these acts of A to be comprehensible for B, there must be available relevant context, which could, for instance, be provided by previous utterances. A typical type of context for these communicative acts would be provided by a question of B, e.g. *What is the proper boxing stance*? *What does your brother look like*?

Continuing in this perspective, we can show that pictures sometimes not only play a similar role as sentences but that they can also play a role similar to paragraphs. A related idea was used by Toulmin and Baier to introduce the function of descriptions as follows:

> Roughly speaking, we can say that John Doe is describing something to Richard Roe only if the purpose of his utterance is of a kind which a picture might serve (Toulmin/Baier 1952, 19).

The upshot of these reflections is that we can assume a remarkable flexibility of the use of pictures as means of communication, which is one of the preconditions of the functional flexibility of text/image combinations. This way of looking at things has the further advantage that we do not have to worry about the "polysemy" of pictures, a problem that I shall not go into here.

As a next step in exploring the function of text/image combinations we have to see how text and image relate to each other as contributions to a total text/image document. In a first, simplifying approach we can distinguish the following types of functional relations between text and image, which could, for instance, occur in science presentations:

(vii) A can describe an object and then *illustrate the description* by showing a picture of the object.
(viii) A can show a picture of an object and then *explain the picture* by drawing attention to features of the picture.

Describing these simple cases in this way we can identify two directions of the functional linkage of text and image, which is related to the sequence of "text and then image" or: "image and then text". In (vii) we have a complex speech act, the description of an object, which is then "supported" by showing a picture. In (viii), on the other hand, we first have a picture being shown, and then this act is "supported" by a complex speech act explaining the picture. The verb *support* here indicates a general type of act, which would have to be specified in more detailed descriptions, e.g., by using action verbs like *illustrate*, *explain*, *document*, etc.

What is missing in these descriptions of patterns is the superordinate function of the text/image combination. Using again a very simple example, we can describe a complex action type as follows:

(ix) A can *represent* an object *by* describing the object *and then* showing a picture of this object.

Here the expression *representing* is again a fairly general description that does not take into account the specific purpose or mode of representation, an aspect that is frequently decisive for choosing a text/image combination. For a more specific description one could, for instance, use a further-reaching *by*-relation like the following:

(x) A can *visualise* an object *by* representing the object *by* describing the object *and then* showing the object.

Or one could use a specification like *vivid* or *intelligible* (*A can represent an object vividly*). In some cases, one will have to describe the use of a multimodal document as a quite complex combination of functional elements, so that it might be a good idea to be flexible in the language of action description, a flexibility that is already present in our everyday language.

The next – and maybe most important – variant of a communicative action by means of text and image is the *simultaneous* use of text and image, for instance the following pattern:

(xi) A can *represent* an object *by* describing the object *and simultaneously* showing a picture of this object.

Another variant of this kind of pattern is a narrative using a text-image constellation:

(xii) A can *tell a story by* describing a series of scenes *and simultaneously* showing a series of images representing these scenes.

These kinds of pattern, which we find in many types of text/image documents, raise the question of how such an arrangement of simultaneously

presented elements functions if there is no predefined direction of functional linkage like in the patterns (vii) and (viii) mentioned above. This question leads to a more sophisticated picture of the *interaction* of text and image as a combined means of communication. A useful starting point for an understanding of this interaction consists in asking how the reception of such documents is to be analysed. This topic will be dealt with in section 6.2, where I shall draw on recent reception studies using eye-tracking technology.

2.7 Discourse types (text types and dialogue types)

With the closing remarks on the Wikipedia article in section 2.5 we already moved to a higher hierarchical level of the structure of texts, namely the *text type*. Text types or genres, as some people prefer to say, are basic structures of whole texts, which include, among other things, the functional structure (speech act types and speech act sequencing patterns), topic structure, knowledge building structures, and the utterance forms of texts. As such they provide criteria for the production and assessment of the global coherence of texts. Examples of text types with quite a variety of textual features are: news articles, crime fiction, scholarly reviews, medical package leaflets, cooking recipes, penalty charge notices, and weather forecasts. A more detailed account of text types and their role for text coherence will be given in sections 4.11 and 5.3, and various text types will be discussed in the course of my exploration of aspects of discourse coherence.

Like in the case of texts we also find *types of dialogues*, which are similarly characterised by basic structures, namely salient speech act types, typical sequencing patterns, topic structures, forms of knowledge management and utterance forms. Examples of such types are planning dialogues, counseling sessions, oral debates, quarrels, narratives, gossip, teaching dialogues, sales talk, examinations, and interviews. In many of these dialogue types we find central speech act patterns and their extensions in sequence patterns, e.g., questions and answers, proposals and objections, advice, claims, arguments and counterarguments, explanations, accusations and excuses, and insults and counter insults.

Finally, there are two complications worth mentioning. Discourses – both texts and dialogues – frequently combine different text types/dialogue types by transitioning from one type to another, for instance when a sober exchange of arguments morphs into a heated quarrel.

The second complication resides in the fact that monological and dialogical discourses are not only related by the use of fundamental communicative

resources but that they also appear together in various forms of combination: Narratives frequently contain reported dialogues, conversations sometimes are *about* texts that are presented in the course of the dialogue, and dialogues (e.g., in social media) consist of small texts strung together in sequences. Thus, there are many good reasons for discussing text and dialogue together.

3. Common ground, context change, and knowledge management

3.1 The hidden part of the communicative iceberg

Discourses are, among other functions, instruments of knowledge generation, knowledge organisation, and knowledge distribution. It is therefore not surprising that knowledge and its dynamics should play a fundamental role in discourse production and comprehension and therefore also for discourse coherence.

In part, knowledge is explicitly communicated in discourse, as in the case of someone explicitly informing someone about a state of affairs by uttering *It's starting to rain again*. But to a large extent, as opposed to the written sentences and the spoken utterances that make up a text or a conversation, knowledge is *implicitly* used and accumulated.[1] It is, so to speak, the hidden part of the communicative iceberg. I hasten to add, however, that salient elements of such knowledge, for instance implicit commitments, may be made explicit, if considered necessary by the participants of a discourse.

The constellation of knowledge most relevant to coherence and understanding in discourse is the knowledge *shared* by the participants of a discourse. This kind of knowledge, which will be defined more precisely in the course of this chapter, has been called *common ground*, a term I shall also be using.

A special case are taken-for-granted "assumptions" like that the surface of a table is solid or that humans (normally) have fingers. For discussions of such background "assumptions", see Wittgenstein's reflections on our accepted "world-picture" ("Weltbild") in "On certainty" (Wittgenstein 1969, §§ 94, 162, 167) and Searle's discussion of the "background", including "preintentional assumptions and presuppositions" (Searle 1983, 154) that form the foundation of our actions.

This chapter will deal with some of the ways in which common ground and knowledge management are involved in communication and thereby contribute to the coherence of discourse.

---

[1] The explicit/implicit divide is not so clear as one might wish, but for the purposes of this chapter the difference assumed should be sufficiently clear.

3.2 Commitment stores, common knowledge, common ground and context

An important part of the implicit knowledge emerging in the course of a discourse consists in the commitments implicitly undertaken by the participants as an aspect of their communicative acts. In the present section I shall elaborate in more detail how commitments come into being, how they evolve and how they become part of the common ground.

As I mentioned in the second chapter, the commitments of speaker/writers become integrated into their commitment store and contribute to both the *options* available for further acts and the *restrictions* on acceptable acts. Using the example of the act pattern of "accusing" mentioned before, I will now show some of the effects of the development of the commitment store, starting with a description of the act pattern:

(1) A can accuse B of having x-ed by asserting that B x-ed by uttering "B x-ed".

In doing so, A undertakes the following commitments:

(i) A believes that B has in fact x-ed,
(ii) A assumes that A can provide evidence for B's x-ing,
(iii) A assumes that B should not have x-ed and that B knows this,
(iv) A assumes that B is responsible and could have acted otherwise.

The first thing worth noting here is that the accusation changes the context in this discourse by placing this battery of commitments in the commitment store of A. This has the consequence of restricting the choice of possible second acts for A. Thus, the following sequences would not be acceptable or, in any case, somewhat odd, provided that the initial act is interpreted as an accusation:

(2) A: (a) You smashed the fruit bowl. (b) *But I don't believe it.
(3) A: (a) You smashed the fruit bowl. (b) ?I didn't see you do it.
(4) A: (a) You smashed the fruit bowl. (b) ?It doesn't matter.
(5) A: (a) You smashed the fruit bowl. (b) ?I know you were pushed from behind.

In each of these cases a commitment of the accusation is either contradicted, retracted or weakened. (2) is a case of a "pragmatic contradiction", where the second act seems to cancel the commitment of the first. This type of contradiction, also called "Moore's paradox", can be used as a heuristic to determine commitments: If the second act of a sequence produces such a contra-

diction, the proposition expressed is the negation of a commitment of the initial act.[2]

With (3b) A admits that a certain strong type of evidence is not available. (But of course, they could have got the information second hand.) One would probably expect A to mark this as a concession and indicate the relative weakness of their position by prefacing their statement with *but*.

With (4b) A retracts the commitment that B's act is bad and contravening a norm. In this case the second act in the sequence could motivate B (and others) to reinterpret the initial act, weakening the accusation to a mere assertion without a normative background: A only *tells* B that they know or noticed that A smashed the fruit bowl, withholding judgment as to whether this should be considered a bad thing or not. (Maybe A always hated this ugly fruit bowl.)

Finally, with (5b) A retracts their assumption of B's responsibility, thereby providing an excuse for B. One may wonder what in this case the point of (5a) could be, maybe just to make clear for the record whose mishap it was.

To sum up, these examples show how commitments interact with the following acts in a sequence by restricting or at least influencing the options for the following acts.

As I mentioned before, there are however also other ways in which commitments can influence the subsequent course of communicative events. One of these ways has been shown before, namely the speaker/writer's option to explain or support their initial act by commenting on the commitments. I repeat here an example given earlier on:

(6) (a) You started smoking again. (b) I can smell it on your clothes.

Here the speaker/writer A asserts that they have evidence for B's x-ing. And this kind of move is possible in relation to other commitments as well.

More importantly, some commitments *open up topics* for general discussions, for example:

(i) the question if and how the *agency* of the actor could be proved,
(ii) the problem if the *norm* in question is really accepted or relevant,
(iii) the question of how far the actor can be held *responsible* or if there are mitigating circumstances.[3]

---

[2] I shall return to this type of inconsistency in the chapter on communication principles (section 8.4.).

[3] Some of the relevant aspects of this matter are discussed in Austin's essay "A plea for excuses" (Austin 1970, 175-204).

Such topics can be discussed either by the "accuser" himself/herself or the "defendant". These types of moves are frequently to be found in controversies where one participant accuses the other of some malpractice and the other defends himself/herself against this accusation.[4] This, by the way, is an example of how closely related structures of text and structures of dialogue are.

The following is an example from a classical controversy between the philosopher Thomas Hobbes and the Anglican bishop John Bramhall. In a pamphlet with the title "A defense of true liberty", published in 1655, Bramhall accused Hobbes of arrogantly criticising learned theologians ("doctors of the Church"), the scholastical practice of "making distinctions", and scholastical expressions in general – which Bramhall tended to use and which Hobbes considered mere jargon:

(7) [Bramhall:] And here he falls into another invective against distinctions and scholastical expressions, and the doctors of the Church, who by this means tyrannized over the understanding of other men. What a presumption is this, for one private man. (Bramhall 1655/1844, 128).

To this accusation Hobbes answered in his pamphlet on "Liberty, necessity, and chance", published in the following year:

(8) [Hobbes:] That he may know I am no enemy to intelligible distinctions, I also will use a distinction in the defence of myself against this accusation. I say therefore that some distinctions are *scholastical* only, and some are *scholastical* and *sapiential* also. Against those that are scholastical only, I do and may inveigh. But against those that are scholastical and sapiential also, I do not inveigh. Likewise some doctors of the Church, as Suarez, Johannes à Duns, and their imitators […] did write such things as neither other men nor themselves understood. These I confess I have a little slighted. Other doctors of the Church, as Martin Luther, Philip Melancthon, John Calvin, William Perkins, and others, that did write their sense clearly, I never slighted, but always very much reverenced and admired. Wherein, then, lies my presumption? If it be because I am a private man, let the Bishop also take heed he contradict not some of those whom the world worthily esteems, lest he also (for he is a private man) be taxed of presumption (Hobbes 1656/1962, 265f.; emphasis added).

Hobbes subtly defends himself by making a distinction himself between mere scholastical expressions and such that are also sapiential, i.e., characterised by wisdom. (This in itself is a polemical move.) As for the purely scholastical expressions, he attacks the norm that he should not polemicize against them ("I do and *may* inveigh"). He then admits that he had indeed

[4] On strategies for answering accusations in controversies, see Fritz (2005).

treated with some disrespect authors who wrote incomprehensibly as opposed to those like Luther and other Protestant theologians who wrote "their sense clearly". In seemingly conceding the validity of part of Bramhall's accusation Hobbes in fact takes the opportunity to renew his original criticism. In a final tit-for-tat response, he reminds Bramhall of the fact that he, Bramhall, is also only "a private man" who could be criticised for his arrogance ("presumption"). In his defence, Hobbes takes up topics provided by Bramhalls's accusations and uses them to support his own standpoint. In this way he uses the structure of the accusation as a resource for the structure and coherence of his own answer. In his response to Bramhall's accusation, he shows that he is very well aware of the commitments his opponent incurred by making his accusation. Thus, elements of the hidden part of the communicative iceberg are brought to light.

This observation leads us to the next general feature of the status of commitments. After A has performed a communicative act and B has understood it, all the aspects of this act, including the commitments of A, become part of B's knowledge. This seemingly trivial fact is essential for the further progress of the communication. The commitments of A not only play an important role for A's own options in his subsequent text, as I mentioned before, but *assumptions* concerning these commitments also determine how B will understand A's initial utterance and further utterances in the text (or dialogue).[5] Let us assume that in the case of the fruit bowl B assumes that A is convinced that it is a good thing that the ugly fruit bowl is finally gone. In this case B will understand A's utterance as a form of praise, not as an accusation. And understanding A's initial act in this way will of course influence B's understanding of A's following utterances in the discourse, some of which may be quite distant in the text.

The next step in our reconstruction of the *context* of A's utterance consists in assuming that after their first utterance A is justified in assuming that their listener/reader B knows A's relevant commitments, which I listed in my description of the act pattern (1), provided, of course, that B understood the utterance correctly. And this again justifies B's assumption that A assumes that B knows A's commitments. So, what we get is a structure of mutually recognised knowledge/assumptions of the participants of the following kind:[6]

---

[5] Readers' and listeners' remarkable capacity for recognising and hypothesising the presence of commitments and *attributing* them to the speaker/writer is a topic in its own right, which I will, however, not go into here.

[6] As the long discussion about common knowledge has shown, this structure is really more complicated, but for my present purpose this model is sufficient.

(9) After A's accusation
(i) A knows that A is committed to (1)(i) to (1)(iv),
(ii) B assumes that A is committed to (1)(i) to (iv),
(iii) A assumes that B assumes that A is committed to (1)(i) to (1)(iv),
(iv) B assumes that A assumes that B assumes that A is committed to (1)(i) to (1)(iv).

Such a structure has been called *mutual knowledge* or *common knowledge*.[7] A related concept is Stalnaker's *common ground*: "It is common ground that φ in a group if all members *accept* (for the purpose of the conversation) that φ, and all *believe* that all accept φ and all *believe* that all *believe* that all accept φ" (Stalnaker 2002, 716).[8] The term *accept* is here used for a family of propositional attitudes that includes belief and assumptions.[9] The latter is the term I used in (9). For the purpose of modelling dialogical communications, we also need a representation of the commitment score of B (or the other participants), which emphasizes the *perspectival character* of this structure.

Common knowledge of the commitments of A (and B) as it develops in the course of the communication is only part of the common ground used by the participants (e.g., the speaker/writer and their listeners/readers). In many cases there are also other bodies of common ground available to the participants, including common knowledge having emerged in their preceding communication history and typical everyday knowledge that can be taken for granted in a given community or society. Not all these bodies of information are available or in the focus of the participants of a communication all the time. Depending on the purpose and the stage of the communication, the stock of common ground actually in focus continually changes. Thus, for every utterance we can assume a certain amount of common ground that is "active" at this point. It is this "active" common ground that we will call the *context* of this utterance.[10]

Every communicative act presupposes a context and also *changes* the context. In so doing every act provides resources for further acts. Thus, common ground and the context are both products and resources of the communication at hand. And tracking the flow of context change is one of the central mechanisms of coherence management, both for speaker/writer and listener/reader.

---

7 The term *mutual knowledge* is used by Schiffer (1972), among others, *common knowledge* by Lewis (1969) and many others.

8 For the concept of common ground and the procedures of creating common ground ("grounding"), see also Clark (1996, ch. 4 and 8).

9 For Stalnaker's definition of acceptance, see Stalnaker (2002, 716).

10 See Stalnaker (1998).

3.3 The dynamics of commitment stores and context change

The observations on the status of common knowledge, context and context change made in the last section will now be used to discuss some examples of the dynamics of commitment stores and their influence on context change.

With every step in the communication, the commitment store of the speaker/writer changes incrementally, and as the audience follows the utterances of the speaker/writer, this changing commitment score is also incorporated in the common knowledge of the participants. A simple example of a short text will show how this mechanism works. Let us take the following sequence of utterances from the beginning of a paper (Tomasello 2006, 506):

(10) (a) **Why don't apes point?** (b) Chimpanzees gesture to one another regularly. (c) Although some of their gestures are relatively inflexible displays invariably elicited by particular environmental events, (d) an important subset are learned by individuals and used flexibly – (e) such things as "arm raise" to elicit play or "touch side" to request nursing. (f) We know that such gestures are learned (g) because in many cases only some individuals use them […]. (h) And their flexible use has been repeatedly documented […].

In the following pages, I shall sketch an analysis of this piece of text in three steps, by first naming the kinds of acts performed with the individual utterances, then focusing on the stepwise addition of commitments (including some comments on the formulation of the commitments), and, finally, giving a survey of the incremental process of the emergence of the relevant body of commitments. This procedure also models aspects of the reader's comprehension of the text. I begin by presenting the sequence of speech acts performed by the sentences in (10).

The sequence of speech acts performed by sentences in (10):

(10a) *announces* the topic of the paper.

(10b) is *an introduction* by means of the *general assertion* that chimpanzees gesture.

(10c) is the (*partial*) *concession* that some gestures are (only) inflexible displays invariably elicited by particular environmental events.

(10d) is the *contrastive assertion* (contrasting with (10c)) that an important subset (of gestures) are learned by individuals and used flexibly.

(10e) *gives examples* for the subset of gestures mentioned in (10d).

(10f) *supports the assertion* (10d) *by asserting that it is known* (by experts) that some gestures are learned.

(10g) *justifies the assertion* (10f) *by giving a reason* why it is legitimate to assume knowledge of the relevant fact.

(10h) *supports the assertion (*10f) by asserting (in addition) that flexible use of gestures is well-documented.

The list on the following two pages shows the commitments undertaken with each speech act of text (10) and also some comments on these commitments. For convenience the utterances of (10) are repeated. (The following abbreviations are used: “U” = utterance, “SP” = speech act description, “CO” = commitment(s)).

(10a)

U *Why don't apes point?*

SP (10a) *announces the topic* of the paper.

CO (i) A accepts that apes don't point.
(ii) A is prepared to explain why apes don't point.

(10b)

U *Chimpanzees gesture to one another regularly*.

SP (10b) is *an introduction* by means of the *general assertion* that chimpanzees gesture.

CO (iii) A accepts that chimpanzees gesture.
(iv) A assumes that this proposition is uncontroversial.

Comment: At this point (iv) is a hypothesis of the reader, which is however confirmed by the fact that the author does not attempt to support this proposition later on in the text.

(10c)

U *Although some of their gestures are relatively inflexible displays invariably elicited by particular environmental events*

SP (10c) is the (*partial*) *concession* that some gestures are (only) inflexible displays invariably elicited by particular environmental events.

CO (v) A accepts that *some* gestures by apes are instinctive behaviour.
(vi) A assumes that some people might believe that all gestures by apes are instinctive behaviour.
(vii) A accepts that there are also (other) gestures that are not instinctive.

Comments: (v) Is based on the assumption that "inflexible displays invariably elicited by particular environmental events" can be taken as a definition of instinctive behaviour. (vi) follows from the interpretation of (10c) as a concession. (vii) is based on an implicature that is suggested by the partial concession.

(10d)

U *an important subset are learned by individuals and used flexibly* –

SP (10d) is the *contrastive assertion* (contrasting with (10c)) that an important subset (of gestures) *are* learned by individuals and used flexibly.

CO (viii) A accepts that an important subset (of gestures) are learned by individuals and used flexibly.
(ix) A assumes that these gestures – in contrast to the gestures in (10c) – are not purely instinctive behaviour.
(x) A is prepared to support this assertion (on demand).

(10e)

U *such things as* "*arm raise*" *to elicit play or* "*touch side*" *to request nursing*.

SP (10e) *gives examples* for the subset of gestures mentioned in (10d).

CO (xi) A assumes that it is useful/necessary at this point to illustrate/substantiate his assertion (10d).[11]

Comment: (xi) is based on the implicature suggested by the fact that A gives this illustration.

(10f)

U *We know that such gestures are learned*

SP (10f) *supports the assertion* (10d) *by asserting that it is known* (by experts) that some gestures are learned.

CO (xii) A confirms the commitment (viii).
(xiii) A is prepared to justify the assertion (10f).

Comment: (xii) The proposition asserted with (10f) is not completely new. It is already present in a slightly weaker form in commitment (viii), so (10f) contains an element of repetition.

(10g)

U *because in many cases only some individuals use them*

SP (10g) *justifies the assertion* (10f) *by giving a reason* why it is legitimate to assume knowledge of the relevant fact.

CO (xiv) A assumes that the fact that in many cases only some individuals use these gestures justifies the assertion (10f) that it is known that these gestures are learned.

(10h)

U *and their flexible use has been repeatedly documented*.

SP (10h) *further supports the assertion* (10f) by asserting that the flexible use of gestures is well-documented.

CO (xv) A assumes that the flexible use of gestures is well-documented.

Now the incremental accumulation of commitments can be shown in the following table. Particularly salient commitments, which are also closely connected and form a kind of "take-away" from this passage, are marked in red.

[11] Examples can either be used to *illustrate* an (abstract) assertion or to *substantiate* a (controversial) assertion.

10a	10b	10c	10d	10e	10f	10g	10h
(i)	(i)	(i)	(i)	(i)	(i)	(i)	(i)
(ii)	(ii)	(ii)	(ii)	(ii)	(ii)	(ii)	(ii)
	(iii)	(iii)	(iii)	(iii)	(iii)	(iii)	(iii)
	(iv)	(iv)	(iv)	(iv)	(iv)	(iv)	(iv)
		(v)	(v)	(v)	(v)	(v)	(v)
		(vi)	(vi)	(vi)	(vi)	(vi)	(vi)
		(vii)	(vii)	(vii)	(vii)	(vii)	(vii)
			(viii)	(viii)	(viii)	(viii)	(viii)
			(ix)	(ix)	(ix)	(ix)	(ix)
			(x)	(x)	(x)	(x)	(x)
				(xi)	(xi)	(xi)	(xi)
					(xii)	(xii)	(xii)
					(xiii)	(xiii)	(xiii)
						(xiv)	(xiv)
							(xv)

This table shows how after every utterance the commitment store is updated.[12] We see that after eight utterances quite a considerable commitment store has accumulated, creating a web of commitments. What is essential to mention at this point is that this process of updating can also be assumed for the reader. Thus, after reading (10a) to (10h) the reader has accumulated *knowledge* of these commitments or, to speak more carefully, *assumptions* concerning the writer's commitment store and its dynamics. And as the writer can hope to have produced this knowledge (or these assumptions) in his reader and that the reader will also assume this as given, we can now assume – if nothing went wrong – that the structure of commitments presented in the table is now common ground and that the table therefore also shows (part of) the incremental *change of context*.

Research from psycholinguistics has convincingly shown that the assumption of some sort of updating mechanism active during text comprehension is a useful part of models of text comprehension,[13] so that, in some sense, the updating procedure modelled here can be assumed to have psychological reality. An interesting fact in this connection is that frequently the surface expressions used to perform linguistic acts are not kept in long-term memory.

---

[12] See Thomason (1990, 338): "The goal of asserting a proposition is updating the conversational record with the proposition; and asserting a proposition is the standard way of updating the conversational record".

[13] See, for example, Millis/King/Kim (2000).

The same seems to be true of the total body of commitments, whereas *salient* subtopics and probably *salient* acts and their commitments are more permanently preserved in memory, embedded in larger structures like major topics, major act sequences, major configurations of propositions or event structures (in the case of narratives). The fact that salient commitments are frequently monitored on the basis of their relevance is shown by the practice of making them explicit in cases of "trouble", e.g., in the case where a hearer/reader suspects an (important) inconsistency.

The types of acts performed by (10a) to (10h), which leave their traces in the above body of commitments, are characteristic of academic writing. They include general assertions, specific assertions, a concession, presenting a contrast of two assertions, giving examples, and providing a justification for an assertion (i.e. an argument for this assertion). The central assertion of this part of the paper is (10d) – that some gestures are learned (and not instinctive). It is repeated in a slightly strengthened form by sentence (10f) and justified by means of (10g). The relevant commitments of these acts are marked red in the above table. So what the author has achieved by this point in this text is: He has introduced the topic of gestures, differentiated between two types of gestures, providing criteria for the two types (e.g. being learned by individuals and used flexibly for the second type), and he has emphasised the point that the second type can be legitimately assumed in apes.

For the paper as a whole, this introductory part can be understood as a kind of orientation on the topic of apes' gestures, which can serve as a background to the main topic of pointing as a characteristically human faculty. Moreover, the author has provided the background for a more specific version of his title question: If apes have the ability to gesture flexibly, why do they not use pointing gestures?

As for the *coherence* of this passage, the topic of chimpanzees' gestures introduced by the first generalising sentence (10b), plays a major role.[14] As part of our general knowledge, gestures can be considered a superordinate topic, of which the various types of gestures (including pointing) form subordinate topics, with pointing being the central topic of the paper. I shall have more to say about topics in chapter 4.

The development of the topic of chimpanzees' gestures is implemented in the text by different types of coreference devices: NP with possessive determiner "*their gestures*" (10c), ("*their use*") (10h), elliptically ("*an important*

[14] To see the connection between the headline "Why don't apes point?" and the first assertion the reader will have to use their knowledge that pointing is a kind of gesture.

subset" [of gestures]) (10d), generalising ("*such gestures*") (10f), and anaphoric pronoun ("*them*") (10g).

A second major factor of coherence in this passage is the connection between different acts, both consecutive acts and non-adjacent acts. Closely connected *consecutive acts* are the contrasting pair (10c) and (10d), the exemplification sequence (10e) and (10d), and the sequence "assertion & justification" (10f) and (10g). *Non-adjacent* acts that are connected are the two assertions of the main point (learned gestures of apes) (10d) and (10f) and the assertion (10h) about documentation of flexible gestures, which supports the mention of the second criterion (flexibility) in (10f). For "distant" acts, in particular, tracking the dynamics of the commitment store plays an important role, because it is there that the relationships of the relevant commitments are "documented". For example, the status of a "second assertion" of the same proposition can be "read off" from the state of the commitment store.

As these observation show, the passage discussed here is tightly organised by the application of closely interwoven organising principles, i.e., act-sequencing patterns, commitment management, topic management, and the use of coreference devices. And this practice continues throughout Tomasello's paper, which I will, however, not discuss further.

3.4 Long-distance effects of context change

After having elaborated on some local effects of context change, I will now briefly discuss some long-distance or global effects. One of these effects has already been mentioned, namely that some commitments open up topics to be dealt with later in the text. The second type I shall mention here, which in many cases is closely related to the first, is the well-known option that a newly introduced entity can be referred to again at a later point of the text. To open up this option the requisite context has to be created, as I shall now show.

By introducing an object into the discourse, the speaker/writer incurs the new commitment that this object exists – at least in the relevant universe of discourse –, and this commitment becomes common knowledge. Thus, we have a change of context that is the basis for later reference to this object. Take the following example from the opening passage of Douglas Adams' "Hitch-Hiker's Guide", part of which I already quoted earlier on:

(11) (1)(a) The house stood on a slight rise just on the edge of the village. (b) It stood on its own and looked over a broad spread of West Country farmland. (c) Not a remarkable house by any means – it was about thirty years old, squattish, squarish, made of brick, and had four windows set in the front of a size and proportion which more or less exactly failed to please the eye.
(d) The only person for whom the house was in any way special was Arthur Dent, and that was only because it happened to be the one he lived in. […]
(e) It hadn't properly registered yet with Arthur that the council wanted to knock it down and build a bypass instead (Adams 1979, 9; emphasis added).

The house is introduced by the narrator into the make-believe world by the definite description *the house*, which commits the narrator to the existence of this – as yet unidentified – house in the relevant story world.[15] The standard type of phrase for the introduction of new objects, which is however not used in this case, is *a house*. However, the book begins with a kind of preface, printed in italics, that ends as follows: "… this remarkable story begins very simply. It begins with a house." So, at least in the preface, we find the standard form, which also commits the speaker/writer to the existence of a specific house.

The introduction of the house with its concomitant commitment of existence now permits the narrator to give a *description* of the house, including the identification of its owner. In the course of the description the narrator refers to it again locally by the use of the anaphoric pronoun *it* and the unspecific definite description *the house*. Introducing the owner of the house gives the narrator the further option to refer to the house later on with expressions like *my house* and *your house*. With sentence (11e) he introduces the complication which forms the topic of the next few passages, namely the planned demolition of the house and the attempts of the owner to prevent it. After these passages there is a longer change of scene – one and a half chapters –, where Arthur Dent and his friend Ford Prefect spend some time drinking in a pub and Ford Prefect tries to explain to Arthur Dent that the world is going to end. Suddenly they hear a rumbling crash from outside, and the following dialogue ensues:

(12) (a) 'It's probably just your house being knocked down', said Ford, downing his last pint.
(b) 'What?' shouted Arthur. […] Arthur looked wildly around him and ran to the window.
(c) 'My God they are! They're knocking my house down. What the hell am I doing in the pub, Ford?' (Adams 1979, 26; emphasis added).

[15] For the status of fictional entities, see Searle (1979), Kripke (2011).

At the end of the chapter not only Arthur Dent's house is demolished but also the Earth, which is destroyed by a Vogon spaceship in order to make place for an intergalactic space highway.

What is remarkable from the point of view of reference is that after quite a long textual distance the house can be referred to again without any extra preparations. Once it is introduced and the commitment to its existence is established, referring to it again by appropriate expressions becomes possible. This, of course, is one of the standard procedures of coherence construction.[16]

It is worth noting that the narrator (or a fictional person in the story) could refer to the owner's house any time by using *his house*, *your house* etc., even if it had not been introduced earlier on, thereby obliquely introducing it. But in this case, it would not be the house as described in some detail at the beginning of the story but just the house the owner happens to own.

The principle that commitments stay valid (and alive) throughout a discourse or at least for long distances within a discourse applies to all kinds of commitments, e.g., the commitment to an evaluation of an object. Let us assume that, at the beginning of their discourse, a speaker/writer characterises a certain kind of mushroom as deadly poisonous. If, much later in their following narrative, they report that a person ate a meal of this kind of mushroom, the audience will rightly assume that this person probably died. This inference is based on the speaker/writer's long-distance commitment that they assume these mushrooms to be deadly poisonous.

3.5 Global knowledge management

3.5.1 Knowledge management in narratives

After discussing some aspects of local knowledge management, I shall now proceed to consider questions of *global knowledge management*, which plays an important role for global coherence. This term refers primarily to the strategies a writer has for providing information to his/her readers at the point or points where the readers need it. These strategies largely depend on the type of text or genre the author is writing. Typical examples with specific requirements for "knowledge building" are narratives, game descriptions and

[16] Coreference and the anaphoric use of pronouns are further discussed in section 4.9.

technical manuals. I shall begin by discussing some aspects of knowledge management in narratives.

In order to understand the point of a narrative the hearer/reader usually has to have some information about the protagonists of the story, the place or places where the story plays, the (historical) background of the story and other aspects of the setting. This information can be given at different stages of the narration, e.g., summarised as a single block at the beginning of the narrative or distributed over the whole story, providing knowledge "just in time" at the respective point where it is needed.

A special type of narrative – or rather a family of types – is crime fiction. In this type of text knowledge management is a particularly important issue, and different variants of knowledge management can be used to differentiate different types of crime story. A standard type is the "whodunnit"-narrative, in which the final stage of the story consists in revealing the criminal. In this, as in other types, decisions about when to provide certain bits of information about the crime and the potential perpetrators of the crime play an important role. One of the interesting questions in this context concerns the way in which the beginning of the narrative is to be organised.

An attractive example of what the necessary reflections in this connection are can be found in P.D. James' "Talking about detective fiction" (James 2009), where she discusses, among other things, the question if one should start with the discovery of the body or should rather provide an introductory description of the setting of the story. I quote this passage in full:

(13) Some novelists like to begin either with a murder or with the discovery of the body, an exciting and shocking beginning that not only sets the mood of the novel but involves the reader immediately in drama and action. Although I have used this method with some of my novels, I have more commonly chosen to defer the crime and begin by establishing the setting and by introducing my readers to the victim, the murderer, the suspects, and the life of the community in which the murder will take place. This has the advantage that the setting can be described with more leisure than is practicable once the action is under way, and that many of the facts about the suspects and their possible motives are known and do not have to be revealed at length during the course of the investigation. Deferring the actual murder, apart from the build-up of tension, also ensures that the reader is in possession of more information than is the detective when he arrives at the scene. It is an inviolable rule that the detective should never know more than the reader but there is no injunction against the reader knowing more than the detective […] (James 2009, 54).

In this passage, James presents two different strategies of knowledge management and discusses the pros and cons of either strategy. Generally speaking, one of the secrets of crime stories and the creation of suspense is that certain bits of knowledge are *not* provided early on but are deferred to a later

part of the story. In the words of Elizabeth George, another well-known author of crime fiction: "[…] all suspense actually is the state of wanting to know what's going to happen to the characters and how it's going to happen" and "[…] the type of novels I write require information to be played out with great care. If I give something away to soon, the whole house of cards collapses" (George 2004, 50f.).

These reflections show that decisions about where to give what information largely determine the sequencing of narrative moves. Thus, knowledge management is a fundamental factor for global sequencing.

In sophisticated narrative we sometimes also find a different kind of subtle play with knowledge states, where information given somewhere in the course of the narrative is used much later as a background for the understanding of other passages. Of this kind of strategy, which I shall call "long-distance knowledge management", I shall give two examples.

In the two last paragraphs of Julian Barnes' compact novel „On the sense of an ending" (2011), the protagonist narrates reflections on his attitude towards life. The last four sentences of these reflections, which make up the final paragraph of the novel, go as follows (p. 150):

(14) There is accumulation. There is responsibility. And beyond these, there is unrest. There is great unrest.

These enigmatic sentences begin to acquire a special meaning if the reader searches his knowledge store accumulated in the course of the novel. On p. 5, when relating his schooldays, the narrator presents a dialogue between the history teacher and a rather weak pupil. The teacher asks: "How would you describe Henry VIII's reign?" To which the pupil answers: "There was unrest, sir." (There is an outbreak of barely controlled smirking.) The teacher insists on his question by saying: "Would you, perhaps, care to elaborate?" To this, the pupil replies: "I'd say there was great unrest, sir."

A few pages later, the history teacher asks the pupils to debate the responsibility of Archduke Franz Ferdinand's assassin for starting off the First World War. In the course of this discussion, Adrian Finn, the most intelligent boy of the group, utters doubt concerning "the whole business of ascribing responsibility" and suggests that there was "a chain of individual responsibilities" (p. 12).

Finally, towards the end of the story, after Adrian had committed suicide, the narrator receives a copy of a "fragment" of Adrian's diary, in which Adrian, in the manner of the early Wittgenstein, with numbered paragraphs, reflects on "the question of accumulation" and the question: "to what extent might human relationships be expressed in a mathematical or logical formula". In the course of these reflections, he also asks himself "how far the limits

of responsibility extend" (p. 85f.). And a page later again, the narrator thinks about "the question of accumulation, but not in the sense that Adrian meant, just the simple adding up and adding on life". And these reflections cause in him "a sense of unease, of unrest' (p. 88).

Remembering these earlier passages in the work, the reader realises that the seemingly strange sentences of the ending refer back to important topics of the narrative and may be read as a kind of ambivalent – maybe even ironical – summary of the narrator's thinking on his attitude towards life. From the point of view of the writer, we can see a kind of long-distance knowledge management, which requires a certain amount of attentiveness and memory of the reader.

As I mentioned before, this kind of thing is frequently to be found in sophisticated, tightly constructed literary texts. Further examples could be given from James Joyce. In his study on Joyce's "Ulysses", Stuart Gilbert refers to several snippets in the stream of consciousness of the protagonists, which only become comprehensible to the reader if he remembers earlier episodes in which the relevant object or theme had already been mentioned: "These fragments would seem meaningless to a reader who had forgotten the earlier passages" (Gilbert 1963, 34). This also works in the other direction: Sometimes a puzzling remark is elucidated in a much later passage. "It is for the reader to assemble the fragments and join the images into a band" (Gilbert 1963, 34). Again, this procedure of long-distance knowledge management, which subtly contributes to the coherence of the text, takes a very attentive reader.

3.5.2 Knowledge management in game descriptions

The questions of knowledge management we discussed in respect of narratives also arise in other text types, e.g., in game descriptions, to which I turn now. To understand a game and to be able to play it presupposes certain chunks of knowledge. For card games and board games the following elements of basic knowledge are usually necessary: (i) knowledge about the inventory of cards (in card games) or the inventory of pieces and the board (in board games), (ii) knowledge about the rules (how the cards are to be played, how the pieces can be moved on the board), and (iii) knowledge about good and bad moves, viz. knowledge of strategy.[17]

There is a systematic relationship between these bits of knowledge that

[17] Probably the first information to be given concerns the point of the game or the "game idea", which I shall however not go into here.

has to be taken into account if one wants to produce a plausible sequence of text elements dealing with these bits of knowledge. In general terms one can say: (iii) presupposes (ii) and (ii) presupposes (i). In terms of text elements this means that information on the basic inventory of the game has to be provided first, information on the rules second and information on strategy third. In fact, many game descriptions are structured according to this pattern.

As an example, I quote some passages from Michael Dummett's descriptions of Tarot games:

(15) In its most typical form, the Tarot pack contains 78 cards. There are four suits, of Swords, Batons, Cups and Coins. In each suit there are fourteen cards, the usual ten numeral cards from Ace to 10, and four instead of three court cards, King, Queen, Knight and Jack. In addition, there are twenty-two picture cards, not belonging to any of the four suits. One stands by itself; the Matto, Fool or Excuse (Dummett 1980, 3).

What is interesting here is that, apart from providing information on the pack of cards, (15) also *presupposes* certain things, e.g., the knowledge what a "suit" is and what "the usual numeral cards" are. So Dummett seems to envisage an audience already familiar with some "usual" card games, e.g. Whist or Bridge. But this knowledge can be picked up (or "accommodated") as one goes along. Strictly speaking, however, no information should be presupposed at a given point in the description that was not provided earlier in the text or could not be considered common knowledge for the target audience.

The next important information follows two pages later:

(16) These twenty-one cards [i.e., the picture cards minus the Fool, GF] have a function quite familiar to us: Tarot games are trick-taking games, and the cards numbered from I to XXI serve as permanent trumps.

In imparting the information that the picture cards are permanent trumps Dummett again presupposes some knowledge, namely what a "trick-taking game" is and what "trumps" are. is. With this information the basic inventory of cards has now been introduced and the author can continue by introducing the *rules* of the game. One such rule is the one concerning "following suit":[18]

[18] Explanations of what "taking a trick" and "following suit" means are given in the description of an individual game ("Scarto") later on in the book (Dummett 1980, 14).

(17) As in Whist or Bridge, it is obligatory to follow suit when one can, including playing a trump when trump is led. Unlike those games, if one cannot follow suit when a plain suit is led, it is obligatory to play a trump if one has one (Dummett 1980, 5).

Further information on the rules includes "dealing and discarding", the anti-clockwise play direction and "scoring".

After having introduced the basic rules, the author can now go on to give advice on *strategy*. The following are two examples of the kind of information on strategy the author gives:

(18) It is never wise to lead a King to the first trick, because of the danger that the dealer will trump it. […] Unless the first player wishes himself to capture the Bagatto [i.e. the lowest trump], he will do best to lead from a short suit, not, of course, in such a way as to deprive a Queen or Cavalier of an essential guard.

By using the expressions *it is never wise* and *he will do best* the author evaluates certain options of play and gives advice to the player on what to do and not to do if they want to play successfully. What is interesting about these examples is that, in the first example, the author explains why this move is "never wise" and, in the second example, he describes contexts in which this move is advisable or not. For reasons given above, the topic of strategy is usually placed at the end of a game description.

Although this form of "systematic knowledge building" can be considered a basic strategy of knowledge management, there are also other strategies that can be used in teaching a game. In teaching a board game one could, for example, introduce a certain piece, e.g., the knight in chess, then explain the rule for moving the knight, and finally add the strategic advice that knights are most useful in the centre of the board, because they are most mobile there. In this teaching scenario the complete inventory of pieces and the rules for moving them are not introduced in one go but the combination of rule and strategy is demonstrated by using a single example, which might be more attractive to some learners. But the structure of presupposed knowledge is similar: The presentation of a strategy presupposes the description of a rule, and this again presupposes the introduction of the respective element.

It is interesting to note that instruction manuals, e.g., manuals for household appliances or car manuals, have very much in common with game descriptions. They tend to describe the inventory of buttons and switches, their function – i.e., the 'actions' one can perform by pressing them or turning them on –, and – here there is a difference – possible problems that may occur in using the object and solutions to these problems.

3.6 Common ground management and "information packaging"

When discussing questions of common ground management, we found that speaker/writers tend to take into account the knowledge state of their hearer/readers at a given point in the discourse/text, performing a kind of "recipient design". Now this not only applies to the practice of building up knowledge in a way that accommodates the hearer/reader's incremental accumulation of knowledge but also to the ways speakers formulate their "message" in order to *signal* to the hearer/readers in what way given information – or part of it – is to be taken, e.g., as something that is already in the common ground or as something that is to be taken as relevant new information. This kind of signalling is performed by syntactical and prosodical means, e.g. by so-called information-packaging constructions (cf. Huddleston/Pullum 2002, 1365ff.) or by nuclear intonation/stress, following language-specific rules. Information packaging contributes to a smooth flow of information.

Before going into more detail, I should like to insert a remark on the use of the expression *information packaging*. This metaphor, introduced by Chafe (1976, 28) and used by many authors since, does not really go well with the action-theoretic framework of the present book. In performing linguistic acts, we utter certain expressions with certain intentions and certain propositional contents, assuming that the hearer/reader will "pick up" what we mean on the basis of their knowledge of these expressions and the common ground presupposed by the utterance. Thus, there is really no process of "packaging" information into a given container at all, and the container metaphor seems misplaced. What does happen is that the speaker/writer will use expressions of a form that, they assume, will permit the hearer/reader to understand the speaker/writer's meaning and thereby to acquire the information to be transmitted. Now, in spite of these misgivings with the container metaphor, I shall for the rest of this chapter use the expression *information packaging*, as it is well-established and has a certain lively appeal.

I shall start with an example of a two-sentence sequence, in which a report on recent events is followed by a prognosis of the possible consequences of these events:

(19) (a) The UK and EU have become embroiled in a series of rows over vaccines [...]. (b) These may scar relations for some time to come. (Guardian Online, 02/02/2021, Opinion)[19]

In this case the definite noun phrase "the UK and EU" presents the entities about which the new information is given, whereas "a series of rows over vaccines" presents salient new information. The expression used to express what the statement is *about* is sometimes called the "sentence topic"[20] or "topic constituent". Some languages also have special devices to indicate the topic constituent; in English, for example, there is the *as for* construction:[21]

(20) As for the UK and EU, they have become embroiled in a series of rows over vaccines.

This construction is frequently used to signal a new topic, for example in a context where (20) follows after a discussion of the difficult relations between *other* countries, e.g., the USA and China. Sometimes it is also used to lead back to a topic dealt with earlier in the discourse.

Now let us return to example (19a). As I mentioned, the salient information provided with this statement is presented by constituents later in the sentence. This conforms to a principle of information flow (in English), namely that salient information, the *focus* of a used sentence, is normally presented at the end of the sentence and often marked by sentence accent, whereas the topic constituent is placed in the subject position at the beginning of the sentence. It is possible to isolate the focus constituents by presenting the sentence as an answer to a given question:

(21) (a) A: What have the UK and the EU become involved in? (b) B: (The UK and the EU have become embroiled in) a series of rows over vaccines.

Using this method, the accented focus constituents in our example could now be indicated as follows: "[a series of ROWs over VACcines]$_F$". "Thus, if the information principle is being observed, it is new information in this last element that is highlighted in this way" (Biber et al. 1999, 897). Seen from the point of view of the discourse context, it is this information that is meant to be newly incorporated into the common ground.

Apart from highlighting salient information, focus constituents sometimes also have other functions, including the highlighting of corrections and paral-

---

[19] The headline of the article is: "The vaccine rows are just the first of many spats between the EU and the UK". (https://www.theguardian.com/commentisfree/2021/feb/02/vaccine-rows-spats-eu-uk-competitor). (02.02.2021).

[20] Some problems of the term "sentence topic" will be discussed in section 3.7.

[21] For a discussion of the status of the *as for* construction and its relation to left dislocation, see Huddleston/Pullum (2002, 1409f.).

lels as in the following examples:[22]

(22) (a) It was not MAry who stole the cookies, (b) it was PEter.

In this case it is the combination of the cleft construction ("It was not … who") with the focus marking that serves to mark the contrast that is used for the correction.

(23) (a) MAry stole the COOkie (b) and PEter stole the CHOcolate.

Here the double parallel of the structure is emphasised by the placement of the sentence accent.

Another element of information packaging worth mentioning at this point is the marking of *givenness*. Information that is assumed by the speaker/writer to be already present in the recent or active context of an utterance is sometimes called *discourse given* information and marked as such. A frequent device for signalling discourse givenness is the anaphoric use of pronouns which refer back to an entity or entities recently mentioned in the discourse. An example can be found in our sentences (19) above, which I repeat here for convenience:

(24) (a) The UK and EU have become embroiled in a series of rows over vaccines [...]. (b) *These* may scar relations for some time to come.

After the series of rows over vaccines has been introduced as new information, the writer intends to give some additional information about these events in (24b), and in order to signal that he intends to do so he uses the anaphoric pronoun *these* to refer to these rows. (Later on in the text, at some distance from (24a) he might prefer the noun phrase *these rows* or *the rows mentioned before*.)

Generally speaking, definite noun phrases – or definite descriptions, as they are called in the theory of reference – are often used to identify entities like persons, concrete objects or abstract objects which are assumed to be known to the hearer/reader, either because they were mentioned before in the text or because their identity is part of common knowledge (in a given community). Such entities can, for example, be referred to by *the present Prime Minister*, *the Covid virus*, and *the theory of reference*.

In some text types we find a whole series of anaphoric pronouns signalling that the writer continues to write about a given entity. This is the case in the following extract from "The moving toyshop", a well-known mystery by Edmund Crispin, in which, like in many narratives, *new* persons or objects

[22] See Krifka (2006, 25).

are introduced by using an *indefinite* noun phrase.[23]

After the poet Cadogan, one of the protagonists of the story, finds the corpse of a murdered old lady, he and the amateur detective Professor Gervase Fen are looking for a girl who might know something about the murder. This is the point at which a mysterious girl (and her dog) is introduced in the story. (The referring expressions are highlighted in italics.)

(25) *Cadogan* looked. *A girl* had just emerged from an alley-way which ran behind one of the shops in the Cornmarket. *She* was about twenty-three, tall, with a finely-proportioned, loose-limbed body, naturally golden hair, big candid blue eyes, high cheek bones, and a firmly moulded chin. […]. *She* wore a shirt and tie, a dark brown coat and skirt, and brogue shoes, and walked with the insouciant swinging grace of perfect health. And beside her trotted *a Dalmatian dog*.
[…]. *Cadogan* said: '*That* must be *the girl*.'
She heard, saw them, and stopped. The lingering smile faded from *her red lips*. With something like panic in *her eyes*, *she* changed direction and cut across the road, walking so fast down Broad Street that *she* almost ran, and glancing back over her shoulder. […]. For a moment *the girl* was lost from *their view* behind the windows of a large china shop, but *they* soon caught sight of *her* again, pushing hastily through the ambling crowds on the pavement. By mutual consent *they* began to run after *her*. […]. *The girl*, too, began to run as *they* drew nearer to *her*, *the dog* cantering beside *her*.

After introducing the girl, using the indefinite description *a girl*, the narrator gives a detailed description of the girl, regularly using anaphoric pronouns to identify her as the sentence topic.[24] In addition, he also introduces the dog by using an indefinite description again (*a Dalmation dog*). In his utterance "That must be the girl" Cadogan refers to the girl they are looking for. At this point it is however not clear if the girl they are looking for is identical with the girl Cadogan is now seeing. Reference to the girl described is taken up again by the use of anaphoric personal pronouns (*she*) and noun phrases with possessive pronouns (*her red lips*, *her eyes*). After a short passage where the perspective of the two "detectives" is taken, reference to the girl is taken up again by using the definite description (*the girl*) and similarly after another short passage describing the activities of the "detectives". Finally, reference to the dog is also taken up again by a definite description (*the dog*). The two "detectives" are similarly referred to with *they* and *their view*.

These passages from the novel show in nearly prototypical form how anaphoric pronouns are used to secure continuing reference to a given person (or object) in the course of a detailed description (and the following narration)

---

[23] Crispin (1958, 70f.).

[24] For a discussion of this and other uses of indefinite noun phrases, cf. section 4.10.

and how definite descriptions are also used to re-identify a person (or object) mentioned before in the discourse, in this case, however, at a point of the story where the salient person (the girl) was not mentioned for some sentences. In this example givenness coincides with sentence topic as the anaphoric pronouns are placed in sentence topic position. This is one of the reasons why in earlier discussions givenness and sentence topic were sometimes confused. But the given entity can also be presented in other syntactical positions, like in the following example:

(26) (a) *The girl* crossed the street. (b) Two men in dark coats followed *her*. (c) When *she* saw *them*, *she* was frightened and ran away.

In some cases, the continuity of the givenness indicators, e.g., anaphoric pronouns, in sentence topic position is secured by an additional information-packaging construction, namely a passive construction as in the following example:

(27) (a) *The girl* crossed the street. (b) *She* was followed by two men in dark coats. (c) When *she* saw them, *she* was frightened and ran away.

Using the passive voice in this way to ensure the continuity of the syntactic structure contributes to the even flow of information.

As a final observation on questions of information packaging I shall give a few examples for problems of "information density", which frequently arise in the context of translation. I shall start from a paper on "information packaging and translation" by Cathrine Fabricius-Hansen, where she noted that "by simply exploiting the structural possibility of expanding the noun phrase to the left and to the right at the same time, one may pack more information into a German noun phrase than could be handled, in a reasonably balanced and unambiguous way, within an English or a Norwegian noun phrase" (Fabricius-Hansen 1999, 203f.). This concerns in particular the use of extended adjectival or participal pre-nuclear modification of noun phrases. To illustrate this phenomenon, I use an example from a recent book on the history of the Early Medieval migration period ("Völkerwanderung") by a German historian (Meier 2019, 482). (The relevant structural properties of the sentence, the head noun and the pre-nuclear participle construction, are marked by underlining and italics.)

(28) Die älteren, *von ethnischen Interpretationsmustern geprägten* Ansätze, die im Massaker des Jahres 471 den Höhepunkt eines Machtkampfes zwischen ‚Germanen' und ‚Isauriern' sahen, sind inzwischen widerlegt.

When rendering this sentence in English, the translator has to deal with the "high-density" participle construction left of the head noun (*Ansätze*), which has no direct counterpart in English syntax. The standard solution to this

problem consists in rendering the participle construction as a relative clause or a post-nuclear participle construction:

(29) The earlier (historical) approaches, (which were) characterised by ethnical patterns of interpretation

In choosing this solution, the translator fills the slot (right of the head noun) that is taken up by the long relative clause ("die … sahen") in the German version. He therefore has to find a different slot for the translation of this clause. This problem can be solved by rendering the relative clause as a verbal phrase, thereby closing the sentence:

(30) The earlier (historical) approaches, (which were) characterised by ethnical patterns of interpretation, considered the massacre of 471 as the culmination of a power struggle between "Germanic" tribes and "Isaurians".

Now there is still an important bit of information from (28) missing, namely the statement that these approaches have in the meantime been refuted. The thing to do is to use what Fabricius-Hansen calls a "sentence splitting strategy", an information-packing strategy which consists in presenting this information in a sentence of its own (Fabricius-Hansen 1999, 186):

(31) The earlier (historical) approaches, characterised by ethnical patterns of interpretation, considered the massacre of 471 as the culmination of a power struggle between "Germanic" tribes and "Isaurians". These approaches have in the meantime been refuted.

By using this strategy, the information is split into two incremental chunks which are connected by the coreferential noun phrase "these approaches". As a result, all the information presented in (28) is rendered, however with a difference in the "weighting" of the information. The original relative clause is "upgraded" to a verbal phrase that contains the main focus of the first sentence, whereas the original main focus is transferred to a sentence of its own. In this way "information density" is reduced, which may improve comprehensibility, but the original balance of the information weights is given up. There is also a second solution, which follows the sentence splitting strategy as well, however with a different choice of splitting:

(32) The earlier (historical) approaches, (which were) characterised by ethnical patterns of interpretation, have in the meantime been refuted. They considered the massacre of 471 as the culmination of a power struggle between "Germanic" tribes and "Isaurians".

In this solution the original main sentence with its focus is preserved and the information is also all there, but again the balance of the information weight is changed. In this case, the "upgrading" of the relative clause to a long sentence of its own seems to give the information provided more weight than

originally intended with the relative clause. Thus, there seems to be an information-packaging dilemma that is hard to avoid by the translator. What this example also shows is a way in which information packaging interacts with the creation of coherence.

3.7 Sentence topic and discourse topic

Reading the preceding section, the reader might have noticed that the term *topic* was used there in a different sense from the way it was used in other places of the present book – and will be used in the chapter on topic management. In order to forestall possible confusions, I shall now briefly discuss these two uses of *topic*. In everyday parlance the topic of a conversation or a text is what this discourse is *about*. In conversations, discourse topics sometimes *emerge*, sometimes they are predefined, as for example in an examination. As for texts, it is frequently the case that the writer sets himself a topic – or is given a topic – on which he then intends to elaborate. But in the writing process, too, the topic may develop in the course of the work. In any case, the concept of topic used in this context always relates to a sequence of sentences (or speech acts), e.g., a paragraph, a sequence of paragraphs, or a whole text/discourse. It provides relevance criteria for individual utterances belonging to this sequence. In some cases, the topic of a discourse is explicitly formulated, but this is not necessarily the case. For example, a (passage of) text may be *about* environmental problems, without the expression *environmental problems* occurring in the text.[25]

In contrast, a *sentence topic* – or a *topic constituent*, as it is also called – is a part of an individual sentence that serves to indicate under which entity the new information presented with this sentence is to be entered into the common ground "file", in other words, to indicate what the respective speech act – e.g. a statement – is about. This kind of "aboutness" concerns the identification of a certain entity, whereas the aboutness of discourse topics concerns criteria of relevance for the development of a discourse. For the theory of common ground management, it is unfortunate that these two uses of *topic* have become established, because they tend to create confusion.

The main problem seems to be that there is in fact a possible relationship between these two types of topic, which is, however, rather tenuous. I shall now try to elucidate this relationship. In some texts, like in obituaries and

[25] The concept of discourse topic will be discussed in more detail in chapter 4.

other biographical writings, we frequently find a chain of topic constituents referring to a given entity, namely the protagonist.[26] Each member of this chain is a sentence topic in the sense explained above, but the continuous series of references to this entity can be a reflection of the fact that the whole passage of text is *about* this entity, in other words: that the entity is the *discourse topic* of this text. Thus, it would make sense to say that the (discourse) topic of the obituary is the deceased. In this kind of case there is a connection between a chain of sentence topics and the discourse topic. But such a connection is not necessary for a discourse topic. In the case of the text on environmental problems mentioned above there may not be one single referring use of the expression *environmental problems,* much less a chain of references (in sentence topic position) to this entity. There may, however, be references to air pollution, soil contamination, litter in the ocean, and other things, to which we summarily refer to as environmental problems. Thus, in this case, the relationship between sentence topics and the discourse topic of the text is very indirect, if present at all.

3.8 Making explicit one's knowledge management

In some cases, especially in scholarly writing, authors tend to make explicit aspects of their knowledge management, e.g., in cases of repetitions, like in the following example:[27]

(33) Several of the terms in [4] are based on processes, and *it is worth repeating* at the beginning of this chapter that process terms are to be interpreted simply as convenient metaphorical means of describing how one construction differs from a syntactically more basic one (cf. Ch. 2, §2). (Cambridge Grammar 2002, 1367).

In this case the author points out the repetition, indicating that it is useful to repeat this comment and even referring to the place where the question was originally discussed. In scholarly writing this kind of reflective move serves to show that the author is making a repetition on purpose and not only out of forgetfulness. At the same time, it can give the reader an idea why it is useful or necessary to make this repetition at this point in the text.

---

[26] For an example of an obituary (Sean Connery), see section 4.9.

[27] The terms referred to in this passage are expressions like *preposing* or *extraposition*. Here and in the following examples the emphasis is added.

(34) *To repeat the conclusions* arrived at: social construction per se does not necessarily undermine truth, […] (Collins 1998, 879).

(35) […] *we have already noted* that certain objects, of a kind most naturally called 'abstract', cannot be considered as possible objects of ostension (Dummett 1973, 489).

Another frequent move serving to make explicit the author's knowledge management consists in explicitly mentioning the knowledge presupposed at a given point of the discourse. Authors frequently introduce such knowledge by prefixing expressions like *here one must know* or *here one must remember*, as in the following example.

(36) So it is not surprising to find claims in ancient authors that some works of art were so perfect that they could not be distinguished from the real thing (*here one must remember* that ancient statues were also painted).
(Jones 2007, 222).

A similar move is made in the following passage, where the authors *remind* the readers of information introduced earlier in the book:

(37) (a) The complementary structure of DNA means that if you have one strand of a double helix you can construct the other. (b) This is because of the nature of DNA and its building blocks – guanine (G), cytosine (C), thymidine (T), and adenine (D). (c) *As a reminder*, whenever there is a G on one strand of the double helix directly across from it on the other strand there will be a C (DeSalle/Yudell 2020, 204).

In fact, both (37b) and (37c) repeat information given earlier in the book, the names of the building blocks of DNA (nucleic acids) in (37b), which are added after the dash, and the explicit reminder about the structure of the double helix in (37c). Both are safety measures to ensure that the reader will understand the following explanations.

A subtle example of this type can be found in Laurence Sterne's *Tristram Shandy*:

(38) I must *remind* the reader, in case he has read the history of King William's wars, – but if he has not, – I then *inform* him that one of the most memorable attacks in that seige [sic] was that which was made by the English and Dutch […] (Sterne 1962, 69; emphasis added).

The author here assumes two types of readers with different backgrounds of knowledge, who he can either *remind* of the attack or *inform* of the attack. He certainly does this tongue in cheek, because probably none of his readers will have read the history of King William's wars and because he is going to tell the story anyway.

There are, of course, also other means of indicating knowledge management, e.g., small print, marginalia, appendices, links in hypertext, and foot-

notes, of which I only want to discuss the latter, at least in passing. The footnote is an interesting device with many potential functions, one of which consists in indicating that the information given there is "extra" information, maybe of secondary importance, which, however, the author does not want to omit altogether. At the same time, the footnote is a means of topic management, as it allows the author to avoid interrupting the topic flow in the main text by presenting extra information externally. The uses and techniques of footnote writing differ from one scholarly field to the other. Historians, for instance, tend to use more and longer footnotes than linguists, and scientists tend to avoid them altogether. Looking for an instructive and at the same time amusing introduction to the history of the footnote, the reader cannot do better than to read Anthony Grafton's small book "The Footnote" (Grafton 1997).

4. Coherence and speech act sequencing

4.1 Sequencing patterns: two-act sequences

Communicative acts do, as a rule, not occur in isolation but rather in combination with others: A question is followed by an answer, and making a claim is often followed by giving a reason for making this claim. As these examples show, in uttering sequences of this kind speakers frequently follow routines or even rules. These regularities I shall call *sequencing patterns*. Sequencing patterns are one of the fundamental resources of coherence in discourse and, therefore, of coordinated communication. [1] They facilitate text production and text comprehension because they provide the speaker/writer with *options* for how to continue after a given speech act and because they provide the hearer/reader with *expectations* concerning the further pathway of the discourse. However, being based on a sequencing pattern is not a necessary condition for a sequence of utterances to be acceptable and comprehensible. It is part of human creativity that speaker/writers also hit upon new sequences that have never before been used (in this way). And hearer/readers are frequently perfectly able and willing to make sense of such new sequences by using various types of inferences and, in so doing, create a coherent understanding of the text for themselves.

One could of course ask why we have act sequences at all. One answer would be that they are just a fact of our natural history. Most forms of human action occur in certain contexts and are connected in sequences of acts. But there is more to be said about it than that. Seen from the point of view of the tasks to be performed in communication it is clear that many communicative aims are more easily achieved by a series of acts than by an isolated act. This is true, for example, of attempts to convince somebody of something, to inform somebody of a complex state of affairs or to coordinate a complex shared activity. In such cases sequences of communicative acts obviously are a useful strategy for successful communication, which consists in making explicit the context and background of salient acts by disclosing the intentions of the actor, the reasons for his act(s), the situation as seen by the speaker/writer, and by referring to the relevant common ground of knowledge. Part of this strategy can be seen as the attempt of the speak-

[1] In the present chapter I will (mainly) discuss monological sequencing patterns, leaving the treatment of dialogical patterns for chapter 13.

er/writer to preempt possible questions and objections concerning this background by providing relevant chunks of information in the immediate context of their respective act. This strategic potential of sequencing patterns may arguably be considered the original reason for the evolutionary discovery and the development of this kind of communicative procedure and the establishment of the practice of connected speech.

Sequences of this kind can be found with requests, demands, refusals, suggestions, recommendations, warnings, and equally assertions, evaluations, accusations, criticisms, and expressions of enthusiasm, just to name a few act types that can be made more comprehensible and acceptable by immediately providing extra background. This strategy of combining an initial act with an immediately following second act is a highly productive pattern for the creation of *two-act sequences* of the following type:

(1) (a) Do not exaggerate your training before the competition. (b) You might get an injury.

The sequence (1)(a)-(b) can be taken as following the sequencing pattern "giving an advice & explaining the advice *by* giving a reason (for the advice)".[2] In this sense, it could be understood as an abbreviated version of the following brief dialogue:

(2) A: Do not exaggerate your training before the competition.
B: Why not?
A: You might get an injury.

Now if, for instance, B tends to doubt the quality or the sincerity of A's advice and A knows this, A could use (1)(b) to *justify* the advice. In this case sequence (1) follows the sequencing pattern "giving an advice & justifying the advice *by* giving a reason (for the advice)", a pattern that is related to the "explaining" pattern but differentiated by the kind of common ground presupposed in its use.

As mentioned before, speaker/writers use sequences of this kind to fulfil the communicative task of providing background, context or justification for their speech acts, thereby "supporting" the initial act, as in example (1). Second acts of this type have a special status in that they have an illocutionary point of their own, for instance as assertions, but they are particularly closely connected to their "primary act".

---

[2] In describing sequencing patterns, I shall use the ampersand (the "&"-symbol), which can be pronounced as "and then", to indicate the sequencing relation, another fundamental action-theoretic relation (cf. Goldman 1970, 21).

4.2 Second acts in a sequence

As I already mentioned in the last section, second acts in sequences may not only be generally related to the act type of the initial act, as in the last examples, but also more specifically to various aspects of the initial act, e.g., the implicit commitments incurred by this act, the proposition expressed by the initial act, and the utterance form of this act.

I shall start by giving an example of the relations of the second act to the commitments of the first, using a simplified description of the act pattern of "evaluating":

(3) A can evaluate an object *by* saying that it is good.
In doing so, A incurs the following commitments (among others):

(i) A assumes that the object fulfils relevant quality criteria for this kind of object.
(ii) A assumes that the evaluation refers to aspects of the object that are relevant for their addressee.

In the case of the evaluation of a bicycle, relevant criteria could be that it is both light and solid, and an aspect relevant for A's addressee B could be its safety. If A assumes that a titanium frame is both light and solid and that disc brakes are particularly safe, and that their addressee knows this, A could utter the following sequence as a positive evaluation:

(4) (a) This is a very good bike. (b) It has a titanium frame and disc brakes.

In this case giving the additional information by uttering (4b) can have the function of an explanation of the evaluation made by (4a). (4b) could be uttered as an answer to the covert question "Why (4a)?". If A is not sure that B knows the relevant properties of titanium frames and disc brakes, A can add this information by extending the sequence:

(5) (a) This is a very good bike. (b) It has a titanium frame and disc brakes. (c) Titanium frames are particularly light and solid and disc brakes are very safe.

(5c) could be seen as an answer to the question "Why is it good for a bike to have a titanium frame and disc brakes?", so it can be described as a kind of backing for the explanation. Example (1) above – giving the reason for an initial-act advice – is another example of this type.

Similar sequences are possible for many initial act types, for example the "accusation" type:

(6) (a) You started smoking again (b) and you know that you should not smoke.

In this case (6b) makes explicit the unspoken norm on which A's accusation (6a) is based – including the assumption that the addressee knows this norm.

Here the explicit assertion of the norm serves to remind the addressee of the norm and/or to intensify the accusation. Another typical commitment of an accusation is the commitment that the speaker/writer has evidence that the addressee did in fact do the misdeed. Relating to this commitment, the speaker/writer could produce the following sequence, which I already presented in the preceding chapter:

(7) (a) You started smoking again. (b) I can smell it on your clothes.

Thus, making explicit an assumption to which the speaker/writer is committed by an initial act can be considered a general pattern for adding a second act in a sequence.

As I mentioned earlier on, basically all aspects of an initial speech act can be referred to in further acts in a sequence. After dealing with the aspect of the commitments implicit in a speech act, I shall now add two more aspects, namely the propositional content and the utterance form of the in initial act.

The following is an example of a second act relating to the propositional content of the initial act:

(8) (a) He handed in the paper in time. (b) It is hard to believe but it is true.

What is hard to believe is the *proposition* that the person handed in the paper in time. Thus, with (8b) the speaker/writer refers to the proposition expressed, using the anaphoric pronoun *it*, and emphasises that the proposition is true. In other cases, it is the *fact* stated in an initial assertion that can be addressed by adding an appropriate second act like in the following example:

(9) (a) Swifts spend their lives in the air, (b) sometimes going ten months without landing. (c) This is a unique fact in the world of birds.

Whereas (9b) can be described as a specification of the fact stated in (9a), (9c) explicitly refers to this fact and emphasises its unique character.

Another frequent second act in a speech act sequence is an explanation of the fact presented by an initial statement, as in the following example:

(10) (a) In recent years, even the best printers have come down significantly in price (b) and this is due to the fact that ink sales are where HP and other printer manufacturers make the majority of their profits.[3]

In this case the nature of the second act as an explanation is made explicit by the use of *this is due to the fact that*, whereas in many other cases the second act must be understood as an explanation on the basis of the sequencing pat-

[3] https://www.techradar.com/news/hp-outlines-plans-to-ensure-printer-ink-stays-lucrative (28.11.2020).

tern and knowledge about the two facts stated in the sequence of utterances. In these cases, the character of explanation remains implicit.

As for comments on the utterance form, various aspects can be referred to in a second speech act, e.g., the syntactic form or the choice of lexical material. In the following example (11a) presents an historical event and (11b) comments on an expression used in this statement ("the technical expression").

(11) (a) These people had been snatched from their communities in Skye and Harris by a representative of their own chief, Norman MacLeod of Dunvegan, to be sold as slaves (or "indentured servants", (b) to use the technical expression).[4]

This kind of comment can also be realised in the form of a parenthesis:

(12) (a) Whenever I post about the #GenderPayGap, people rush to tell me it's a myth. (b) Women "choose" jobs that pay less than men; (c) if they do the same work, they get #equalpay. (d) Both these assertions are, (e) to use the technical expression, bollocks.[5]

In this case, the comment in (12e), which refers to the assertions made by (12b) and (12c), is used ironically, indicating the writer's awareness that she is not really using a "technical expression" but an expression that is "coarse slang", as the OED describes it.

Sometimes also a particular *use* of an expression is explained (e.g., ironical or euphemistic), like in the following example from "The Guardian Today newsletter" online, reporting on a press conference of the British Prime Minister Johnson about talks in Brussels about Brexit:[6]

(13) (a) [Johnson] added on Thursday: "I do think we need to be very, very clear. There is now a strong possibility – a strong possibility – that we will have a solution that is much more like an Australian relationship with the EU than a Canadian relationship with the EU." […] (b) The Australian option is Johnson's name for trading on World Trade Organization terms with the EU, meaning tariffs being imposed on a wide range of goods.

The utterance (13b) here serves to explain to the reader what was meant by the somewhat opaque (and euphemistic) expression *an Australian relationship* in the statement reported in (13a).

---

[4] West Highland Free Press 12.10.07; (https://www.rubh-an-dunain.org.uk/files/RB-The-Dun-Hill.pdf). (09.02.2022).

[5] Catherine Mayer on Twitter (07.01.2018, 10.24) (https://twitter.com/catherine_mayer/status/950115852711530496). (24.06.21).

[6] https://www.theguardian.com/politics/2020/dec/10/boris-johnson-tells-cabinet-to-prepare-for-no-deal-brexit (11.12.2020).

4.3 Relational act types

Some of the second acts mentioned in the preceding sections belong to a group of linguistic acts that play an important role in the organisation of text on account of the fact that their "illocutionary point" consists in contributing to the connection of other acts and the structure of the text. One can, for instance, *contrast* two descriptions, two evaluations, two suggestions etc. This kind of connective function is also characteristic of act types like "supplementation", "continuation", "repetition", "reformulation", "concretising", "precisation", "generalising", "exemplification", "elaboration", "comparing, summarising".[7]

To give an example of the nature of acts of this type, which one could call "relational acts", I shall make a few observations on "elaboration". The point of an elaboration consists in providing extra detail in relation to a statement, a description, a brief narrative, a generalisation, a topic etc. Therefore, the specific character of an elaboration is determined by what it is that is elaborated. An elaboration of a report will present extra details of events, whereas an elaboration of a general statement will provide examples which substantiate the generalisation. On account of this feature, elaborations also contribute to *topic management*. They serve to continue a topic, sometimes specifying it or introducing subtopics. The following is an example of the elaboration of an evaluation that could be taken from an old-fashioned novel:

(14) (a) Mary was a beautiful girl. (b) She had thick, wavy, long black hair and a diamond-shaped face, a smooth, suntanned complexion and high cheek bones. Her eyes were large and deep blue and her slightly arched eyebrows were chestnut brown. [...]

In this case the elaboration consists in a detailed description of aspects of the person which could be considered beauty features. To indicate the continuing topic of the elaboration the writer uses anaphoric pronouns (*she*, *her eyes*).[8]

Generally, acts of these types show *by*-relations of different types: One can insist on an assertion *by* repeating the assertion. One can concretise a general statement *by* giving concrete examples. And one can supplement a description *by* providing an illustration.

---

[7] Some of these nominalisations are rather a mouthful, but they are useful at this point.

[8] The anaphoric use of pronouns is discussed in more detail in the section on reference and coreference (4.9); see also the sections on "information packaging" (3.6) and "constraints" (7.4).

4.4 Multiple options for the second act in a sequence

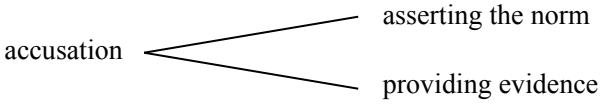
After focusing on simple two-act sequences, I will complement this account by drawing attention to an important complication, viz. the case of multiple options for the second move in a sequence.

Looking at the last examples presented in the last section, one could come to the conclusion that the concept of sequencing pattern used so far is really too restricted. So far, apart from a few three-act sequences, sequencing patterns were all of the form "x & x", e.g., "assertion & assertion". If we go back to our starting point and ask "What can a speaker/writer do after having produced the first sentence (or act) of his text?", we find that in this situation the speaker/writer is confronted by a "textual demand", a task that they have to solve by producing a relevant continuation.[9] Now, in most cases there is not just one solution to this problem, but several options among which the speaker/writer may choose, depending on the particular aim they have in view. And for these options there are also frequently established patterns. So the sequences shown in (6) and (7), repeated here for convenience, can be attributed to the patterns "accusation & asserting the norm" and "accusation & providing evidence".

(15) (a) You started smoking again (b) and you know that you should not smoke.

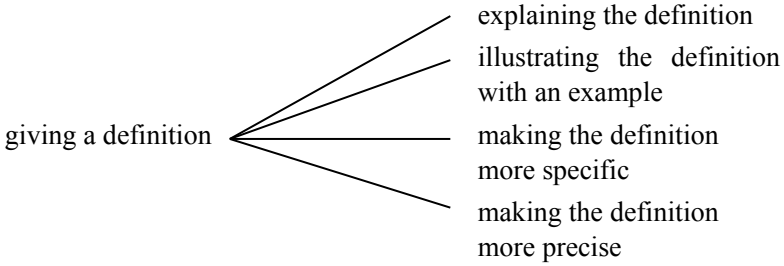
(16) (a) You started smoking again. (b) I can smell it in your clothes.

This sequencing pattern can be presented in the form of an action tree with two edges representing the two alternatives, like in the following example:



In fact, in this and many other cases there are more than two routinised options for the second act in the sequence, which we can represent as an action tree with more than two edges. An example of such a more complex tree is the following representation of the options of second act after "giving a definition" as initial act:

[9] The concept of "textual demand" is derived from the concept of "conversational demand" introduced by Marcelo Dascal in a chapter on "conversational relevance" (Dascal 2003, 37ff.).



To illustrate this tree, I give the following battery of examples:

(17) (a) A text is a sequence of sentences connected by several organising principles. (b) This means that not just any sequence of sentences is to be called a text but only those sequences that conform to certain forms of organisation.

(18) (a) A text is a sequence of sentences connected by several organising principles. (b) One of these organising principles is the topic that structures a text, like in the following example: […].

(19) (a) A text is a sequence of sentences connected by several organising principles, (b) namely sequencing patterns, topics, referential continuity, and incremental knowledge structure.

(20) (a) A text is a sequence of sentences connected by several organising principles. (c) Strictly speaking, it is not sentences but *uses* of sentences that are organised in this way.

In addition to the choice between different options for the second act there is also the option of *cumulating* these options, e.g., "giving a definition & explaining the definition (17b) & illustrating the definition (18b) & making the definition more specific (19b) & making the definition more precise (20b)", which gives the following short paragraph:

(21) (a) A text is a sequence of sentences connected by several organising principles. (b) This means that not just any sequence of sentences is to be called a text but only those sequences that conform to certain forms of organisation. (c) One of these organising principles is the topic that structures a text, like in the following example: […]. (d) [Other organising principles are] sequencing patterns, referential continuity, and incremental knowledge structure. (e) Strictly speaking, it is not sentences but *uses* of sentences that are organised in this way.

This way of looking at things is particularly interesting for the analysis of text production, where the choice of an adequate option and the cumulation of such options is an important strategic aspect, often related to text quality.

4.5 Combinations of two-part sequences and the recursivity of act patterns

A first extension of the patterns presented so far consists in the combination of two (or more) two-part sequences, as in the following example:

(22) "asserting that p & giving evidence for p" + "making the prediction that q & providing backing for the prediction that q"

(23) (a) John is reading the book. (b) He told me so. (c). His friends will soon read it too. (d) They always imitate him.

The whole pattern (22) is made up from two routinely used two-part patterns, but the whole pattern as such is probably not an established routine itself. However, sequence (23), based on this combination of patterns is certainly an acceptable and easily understood bit of text. Therefore, in this case, the speaker/writer can use an extra textual resource to make this chunk of text coherent. The thing to do is to use two sequencing patterns *plus* a developing topic the sequence is about, say "Peter and his friends". The introduction and development of this topic is implemented by the continuity of objects mentioned in (23): John, the book, John's friends. And this continuity is again signalled by the anaphoric use of certain expressions (*he*, *so*, *his*, *it*, *they*, *him*). So what we find here is the use of a combination of two (or three) textual resources for the creation of a coherent sequence of utterances. These extra resources will be dealt with in more detail in section 4.9 (anaphora) and chapter 5 (topics). Generally speaking, such combinations are a most powerful means of text production and account for a good deal of everyday text.

The second clement of complexity to be introduced here is the recursivity of (complex) act patterns. Typical forms of recursivity are arguments within arguments, explanations within explanations, descriptions within descriptions, and narratives within narratives. An interesting pattern in argumentations has the following hierarchical pattern structure:

> "claim & argument$_1$ for claim & argument$_2$ for argument$_1$ & argument$_3$ for argument$_2$".

Each argument can itself be considered a claim, which then has to be given backing by a new argument. This kind of argumentation strategy may become necessary in cases of deep disagreement, where no argument is accepted on its own right.

The same type of structure can also be used for other types of initial act, e.g., an objection to a proposal that is backed by a reason for the objection, which in turn is backed by another reason. This kind of sequence is related to a recursive structure of *why*-questions:

(24) (a) We can't go there tomorrow. (b) I haven't got time. (c) I have to study.

(25) A: We can't go there tomorrow.
B: Why not?
A: I haven't got time.
B: Why?
A: I have to study.

This kind of pattern can be used if one reason is not considered "strong" enough or if the first reason is not self-explanatory.

Further types of complex sequencing patterns, which complement the system of textual building blocks presented so far, will be dealt with in the section on "minimal text types" (section 4.8) and in connection with the discussion of text types (section 4.10.3) and "functional hierarchies" (4.11).

4.6 Sequencing patterns and connectives

In the types of sequencing patterns discussed so far, the utterance form generally contained no special indicator of the kind of relation that held between the two acts. In many cases, however, it is possible to make this relation more explicit by using an indicator. In the case of an explanation of an asserted proposition we have basically two options:

(26) (a) He needs to eat something now. (b) His blood sugar is low.

(27) (a) He needs to eat something now, (b) *because* is blood sugar is low.

In this case the use of *because* makes explicit the character of explanation of (27b), which, however could also be understood without the use of *because*.

It is remarkable that in some languages, certain patterns, for instance the explanation pattern, can be explicitly indicated by quite a variety of markers belonging to different grammatical categories. In German, for example, there are (amongst others) *denn*, *weil*, *da*, *nämlich*, *darum*, *daher*, *weshalb*, *deshalb*, *aus diesem Grunde*, *wegen* used in this function.[10] On the other hand, it is worth noting that some connectives can be used to indicate *different* sequencing patterns, for instance in the case of *since*, which can either indicate a temporal relation or a causal relation, or in the case of *because*, which can either indicate a causal explanation or a justification by giving reasons.

An interesting question is how *frequently* sequencing patterns – or coherence relations – are explicitly signalled by connectives/discourse markers.

[10] For a corpus-based analysis of the discourse functions of causal connectives in German, see Breindl/Walter (2009).

Research on this question has found, first, that this frequency varies considerably with the genre/text type, and secondly, that generally the percentage of explicit markers can be quite low, e.g., 44 % for a corpus of newspaper articles (Taboada 2006, 579).[11]

A type of relation between two acts for which connectives are (almost) necessary to aid understanding are acts of *contrasting*. Take the following example:

(28) (a) In the sentence "A bird hit the car" the phrases *a bird* and *the car* belong to the same category, NP. (b) They have different functions, subject and object, respectively.

Here we have two assertions following each other, but they appear somewhat unconnected. A closer connection becomes apparent if we insert the connective *but* (Cambridge Grammar, 23):

(29) (a) In our example sentence the phrases *a bird* and *the car* belong to the same category, NP, (b) but they have different functions, subject and object, respectively.

The point of (29) is to support the claim that it is necessary for a grammar to differentiate between syntactic categories and syntactic functions. One could think that if two phrases belong to the same category, they also have the same function. This assumption is shown to be unfounded by stating (29b), and the contrast between the original (unfounded) assumption and the correct fact now made explicit is indicated by *but*.

A related kind of contrast can be brought out with the use of *although*. In the following example we might be puzzled about what the assertion (30b) contributes to the sequence, as in this kind of sequence the second sentence frequently contributes an explanation, like in (31b):

(30) (a) Sonia doesn't speak French. (b) She grew up in Paris.

(31) (a) Sonia doesn't speak French. (b) She grew up in Japan.

As the conjunction of the two assertions (30a) and (30b) seems to contradict a common assumption, namely that people growing up in a certain city learn the language spoken in this city, we might suspect some sort of contrast between the assertions (30a) and (30b), but it would be safer for the speaker to make this contrast explicit by using the connective *although*:

(32) (a) Sonia doesn't speak French (b) *although* she grew up in Paris.

---

[11] A similar result was reported for encyclopedia entries (van der Vliet/Redeker 2014, 34).

The Cambridge Grammar explains the contrastive meaning as the implicature – triggered by *although* – that one might expect the superordinate clause (32a) to be false (Cambridge Grammar, 734f.).

Sometimes this contrast is somewhat vague so that it is more difficult to pinpoint the implicature triggered by the connective, as in in the following case:[12]

(33) (a) President Trump inherited 654 miles of border structure along America's 1,900-mile border with Mexico. (b) Over four years, he's constructed 415 miles, (c) *although* of that total, only about 25 miles cover areas that had no previous barriers. (d) The rest replaced or reinforced existing structures.

In this news article with the headline "What will become of Trumps border wall?" the report is meant to show that Trump's promise to build "a big, beautiful wall" from his 2015 campaign was fulfilled only to a relatively small extent. The contrast consists in the small proportion of completely new wall compared to the number of miles only replaced or reinforced. Thus, sentence (33c) expresses a qualification of the statement made in (33b). In cases where the contrast indicated by the connective is vague, it may be necessary to perform an extra interpretation of the relation of the propositions expressed by the two connected sentences.

4.7 Discontinuous sequences

So far, the examples of speech act sequences always represented patterns of adjacent acts. It would, however, be too restrictive to assume that sequencing patterns are always realised in adjacent form.[13] In many cases additional acts are inserted between members of a sequencing pattern, as in the following examples from a chemistry lesson. In the first example we find an assertion (34a) justified by evidence for this assertion (34b):

(34) (a) This solution is alkaline. (b) The red litmus changed to blue.

In the second example, an utterance emphasising the reliability of the assertion ("it is quite clear") is inserted between the two segments of the pattern

[12] The Week, Dec 13, 2020; https://theweek.com/articles/954446/what-become-trumps-border-wall (4.12.2020).

[13] This is an (unrealistic) assumption made by Rhetorical Structure Theory in respect of their "rhetorical relations" connecting spans of text (see section 4.13). Wolf/Gibson (2006, ch. 2) discuss problems involved in representing discontinuous sequences as trees.

"assertion & justification". This kind of strategy might be used if the speaker, for instance a pupil, assumes that their audience might doubt their assertion:

(35) (a) This solution is alkaline. (b) It is quite clear. (c) The red litmus changed to blue.

Frequent insertion elements are glosses, comments, and examples, like in the following three examples (my emphasis):

(36) (a) The first type of example concerns opaque contexts, (b) *as they were called by Quine*, (c) which include sentences with *I believe*.

(37) (a) This strand of RNA, (b) *known as messenger RNA*, *or mRNA*, moves out of the nucleus of a cell to a ribosome, (c) where the genetic sentence is read and translated into a protein (DeSalle/Yudell 2020, 37).

In the case of (37), the additional information is inserted as a parenthesis after the noun phrase *this strand of RNA*.

(38) (a) **Collective nouns** refer to groups of single entities. (b) *Typical examples are*: army, audience, board, […]. (c) All these nouns behave like ordinary countable nouns, (d) i.e. they vary in number and definiteness (Longman Grammar 1999, 247).

(38a) and (38c) present properties of collective nouns, whereas the examples in (38b) are inserted for clarity, and (34d) is added to make more explicit the assertion (38c). This is a well-established sequencing strategy in pedagogical texts.

A type of sequence frequently found with narratives is the insertion of an evaluation into the presentation of an ongoing flow of events, like in the following simple example, which I first present without the inserted evaluation and then with it:[14]

(39) (a) We once had a dog. (b) One day the dog attacked me and bit me in the neck. (c) So my father took him away and shot him (d) because the dog's mother had also been dangerous.

(40) (a) We once had a dog. (b) One day the dog attacked me and bit me in the neck. (x) *That was really bad*. (c) So my father took him away and shot him (d) because the dog's mother had also been dangerous.

Whereas in (39a-c) we have a continuous sequence of event-presenting narrative acts plus an added explanation (39d), in (40) the speaker inserts the evaluation *really bad* (40x) into this sequence, an addition that serves to emphasise the dramatic character of the event. Thus, the inserted evaluation,

[14] This is an abbreviated version of an original story by a ten-year old boy (cf. Fritz 2017, 435).

which produces a discontinuous sequence, makes an important contribution to the meaning of the text.

Inserting extra information into the sequence of narrative acts presenting the main storyline is a well-known strategy, both in everyday storytelling and in literary narrative. The literary technique, called *digressio* (or *egressio*) in classical rhetoric, has also been widely used in various other types of text.[15] Some books, e.g., Robert Burton's "Anatomy of melancholy" (1621) and Laurence Sterne's "Tristram Shandy" (1759-1767), are famous for their digressions, which the authors deliberately introduced and also commented upon. Stern wrote, for instance: "In a word, my work is digressive, and it is progressive too, - and at the same time" (Sterne 1962, 63).[16]

Generally speaking, discontinuous sequences are not uncommon, they are often used strategically in discourse, and in many contexts they can be considered routinised patterns. In the case of complex forms of global discontinuity, writers sometimes explicitly comment on this device to guide their readers, but experienced readers tend to know how to deal with this feature of discourse on their own.

4.8 Minimal text types

There are a number of sequencing patterns that comprise more than two acts in a sequence and that are frequently extended to longer texts by adding or inserting various types of supplementary acts. Such *minimal text types*, as I will call them, include descriptions, narratives and argumentations. In texts of this kind coherence is (partially) based on these sequencing patterns. The following are simple examples of these types.

4.8.1 Descriptions

My first example is *a minimal description* of an edible mushroom:[17]

---

[15] See Quintilian, "Institutio oratoriae", IV.3:15. For a modern discussion of digression and conversational coherence, see Dascal (2003, ch. 10). For an analysis of the function of digressions in a medieval hagiographical text, see Fiona Fritz (2022, ch. 5.3).

[16] For a delightful modern Shandean scholarly work, see Merton (1965).

[17] This description is mainly based on the Wikipedia article (https://en.wikipedia.org/wiki/Boletus_edulis) (26.11.2020).

(41) (a) The Cep or Penny Bun is an edible mushroom. (b) It mainly grows in forests. (c) The fruit body has a large brown cap, (d) and the tubes extending downward from the underside of the cap are whitish, or greenish-yellow, when older. (e) The stem is thick and bulbous and mostly light brown. (f) The Cep is high in protein, vitamins, minerals and dietary fibre.

The description follows the typical sequence found in mushroom field guides. It lists properties of the mushroom, starting with the fact that it is edible and continuing by specifying its habitat (b), its outward appearance – emphasising its colour (c)–(e) –, and its dietary properties (f). This is a typically additive sequence, asserting facts about this kind of mushroom. In addition to the additive sequence of assertions this text is internally organised by topics that are relevant for the practice of mushroom picking: where to find it, how to recognise it, that it is edible and that it has healthy dietary features. Thus, the coherence of this small description is based both on the functional sequencing pattern, i.e., the sequence of acts, and the typical "programme" of topics of such field guides. It is worth noting that such a description can be made more vivid by adding an illustration of the mushroom as an extra aid to identification. In this case description and illustration are combined in a typical pattern of "telling and simultaneously showing".

This minimal descriptive text could be expanded by giving further biological details concerning the spores etc. or by praising the mushroom's taste and giving recipes for cooking it. Thus, the minimal text can form the core of more detailed texts.

4.8.2 Narratives

Similarly, there are *minimal narrative forms*, consisting of an additive sequence of speech acts asserting the occurrence of several related events like the following, which could be triggered by the question "And what did you do today?":

(42) (a) Today, I took the *bus* to the market. (b) I first went around looking for venison, (c) then I bought some vegetables, (d) and finally I had coffee at one of those stalls.

Each of the assertions (42a) – (42d) presents an individual event, with the assumption that the events occurred in exactly the temporal sequence represented by the assertions, what one could call an *iconic* representation. As for the sequence of topics – bus ride, venison, vegetables, coffee –, they are covered by what has been called a typical "frame" for visits to a market. Events of these types are to be expected in a normal, uneventful visit. It would have been surprising and maybe incoherent and hard to understand if the narrator had inserted a sentence like *and then I played a slice*, which could have been expected in a report from a tennis match.

Functionally, the presentation of such a series of events (or actions) can be described as a seriatim application of the pattern of "continuation", the "and then"-pattern. As a standard feature of this pattern, the speaker/writer assumes that the sequence of events is connected by conditional or causal relations of different kinds. Typically, a first action may create conditions for a second action to occur (43) or a second event may be caused by the first (44):

(43) He opened the door and (then) went into the garden.

(44) The dam broke and (then) the valley was flooded.

The narrative core produced by this simple "continuation" pattern can be augmented by various acts of different types, e.g., by giving additional background information (e.g., in a preparatory function), giving explanations, and making evaluations, like in the following extended version of (42):

(45) (a) *It was a really lovely sunny day today*. (b) I took the bus to the market, (c) *because my wife needed the car*. (d) There, I first went around looking for venison, *and I found some very nice deer filet*, (f) then I bought some vegetables (g), and finally I had coffee at one of those stalls.

Thus, as in the case of descriptions, this simple pattern can be extended to form highly complex structures with various types of internal substructures.

On the other hand, there are small snippets of talk about everyday events that also belong to the narrative genre, albeit to a somewhat marginal type. Such forms of event presentation, which are usually embedded in other

communicative or practical activities, have been called "small stories" or "micro narratives".[18] The following is a possible example:

(46) (a) I met Peter in town today, (b) but he was in a hurry.

This remark could, for instance, be made when cooking the evening meal. If Peter is "a person of interest" to the interlocutors, mentioning this event could indicate that there was an opportunity to receive some information about the person's recent activities or his state of health, which was, however missed. If, in addition, it is common knowledge that this person should not have been in town that day, the remark could function as a snippet of gossip. Thus, we see that the genre of narrative comprises a large family of patterns, from very small to highly complex.

4.8.3 Argumentations

Minimal forms of argumentation consist of the sequence "claim & argument for the claim", mostly in the form "making a claim *by* making an assertion & giving an argument for the claim *by* making a second assertion".[19] So without knowing that these assertions are meant as claim and argument we might simply take them for a simple sequence of the type "assertion & assertion", as the following example shows:

(47) (a) The road to human cooperative communication begins with great ape intentional communication, especially as manifest in gestures. […] (b) Apes' learned attention-getters may be the only intentional communicative acts in the nonhuman world that operate with this split-level intentionality: that the other see something and so do something as a result (Tomasello (2008, 320).

In this sequence, the author's evolutionary claim (47a), which might be challenged by some researchers, is supported by the case of the great apes (47b), which uniquely show a kind of intentionality that is also characteristic of human communication. It interesting to note that in the original text the second assertion (47b) – which, with its use of *may*, is somewhat hedged – is only one out of a number of statements supporting the claim (47a). This strategy of cumulating supporting statements is characteristic of Tomasello's "summary of the argument" of his book (Tomasello 2008, section 7.1). In order to understand (47b) as an argument supporting (47a) we must assume

---

[18] For a discussion of forms of "small stories", see Georgakopoulou (2007).

[19] The expression *claim* is used here to refer to an assertion uttered with the commitment that the speaker/writer assumes that the proposition asserted has been or might be challenged (see Toulmin 1958, 97).

that the author is implicitly committed to a proposition of the following kind:[20]

(48) If the apes' attention-getters are a kind of intentional act which is well-known in humans and which is otherwise unique in the nonhuman world, then it is likely that the road to human cooperative communication begins with this kind of great ape intentional communication.

This proposition – or one closely related to it – is not made explicit, but it could be explicitly added to (47a-b) as an answer to the implicit question "In what way does (47b) support (47a)?"

We could extend the category of minimal argumentation by differentiating different types of claim, e.g., statements of fact, conjectures, objections, assessments, accusations, and demands. For the moment, I shall however restrict myself to statements of fact. The next extension consists in a differentiation of different types of arguments in their different functions of supporting the claim. Depending on the kind of claim, there are, for example, legal arguments, statistical arguments, arguments from experience etc. Instead of going into more detail here, I shall now give another example where the basic structure of the minimal argumentation is augmented by other acts, including polemical asides.

In his criticism of "amateur language pontificators", Geoffrey K. Pullum provides an argument against the normative prejudice that one should not use *which* to introduce so-called restrictive relative clauses like *any book which you would want to read*.[21] I shall first reconstruct the minimal structure of the argument and then give an excerpt of the original argument.

The basic claim of Pullums criticism and the supporting argument runs as follows:

(49) (a) To criticise the use of *which* (instead of *that*) in so-called restrictive relative clauses is misguided. (b) This criticism is in conflict with the facts of linguistic usage. (c) The use of *which* is common in excellent writing. (d) [A statistic of the use of *which* in Dickens and other "classical" writers].

The argument (49b) supports the claim made by (49a). In order to function as an argument, the use of (49b) implicitly commits the author to an assertion of the type: *If a linguistic criticism in is conflict with the facts of linguistic usage it is misguided.* (49c) again supports (49b) by substantiating it with a

---

[20] This kind of implicit proposition was called "warrant" by Toulmin (1958, 98) and "inference presupposition" ("Schlusspräsupposition") by Öhlschläger (1979, 88ff.).

[21] This example is taken from a blogpost in the "Language Log" (17.09.04), reprinted in Liberman/Pullum (2006, 36-41). I only use part of the post.

reference to "excellent writing" – as opposed to illiterate writing. And the statistics of (49d) again substantiate the claim made by (49c) by giving empirical evidence. So what we have in (49) is a claim supported by a cascade of arguments.

Now I give the relevant part of the actual text of the blogpost:

> [The critic Sidney Goldberg is] one of those people who believe the old nonsense about *which* being disallowed in what *The Cambridge Grammar* calls integrated relative clauses (the old-fashioned term is "restrictive" or "defining" relative clauses). Strunk and White perpetuate that myth.[22] I've discussed it elsewhere. The notion that phrases like *any book which you would want to read* are ungrammatical is so utterly in conflict with the facts that you can refute it by looking in... well, any book which you would want to read. As I said before about *which* in integrated relatives:
>
> As a check on just how common it is in excellent writing, I searched electronic copies of a few classic novels to find the line on which they first use *which* to introduce an integrated relative with *which*, to tell us how much of the book you would need to read before you ran into an instance:
>
> - A Christmas Carol (Dickens): 1,921 lines, first occurrence on line 217 = 11% of the way through;
> - Alice in Wonderland (Carroll): 1,618 lines, line 143 = 8%;
> - Dracula (Stoker): 9,824 lines, line 8 = less than 1%;
> - Lord Jim (Conrad): 8,045 lines, line 15 = 1%;
> - Moby Dick (Melville): 10,263 lines, line 103 = 1%;
> - Wuthering Heights (Bronte): 7,599 lines, line 56 = 0.736%...
>
> Do I need to go on? No. The point is clear. On average, by the time you've read about 3% of a book by an author who knows how to write you will already have encountered an integrated relative clause beginning with which. They are fully grammatical for everyone. The copy editors are enforcing a rule which has no support at all in the literature that defines what counts as good use of the English language. Their which hunts are pointless time-wasting nonsense.

In this argumentation the minimal argumentation structure is extended by various acts, including a rhetorical question ("Do I need to go on?"), the presentation of statistical material, a repetition of the main argument ("The point is clear. …") and a final polemical flourish ("pointless time-wasting nonsense"), with a word-play ("which hunts"), ridiculing naïve prescrip-

[22] Strunk and White, "The elements of style" is a prescriptive style manual which is a favourite target of Pullum's anti-prescriptivist attacks.

tivists.[23] These extra acts contribute to the polemical and amusing character of this blogpost.

To wind up this section: In naturally occurring texts, which frequently have to serve complex communicative tasks, the minimal text types discussed in this section are often augmented by and hidden behind a complex network of additional acts, which sometimes makes it difficult to locate the core of the minimal structure.

4.9 Speech act sequences, reference and coreference

One of the basic tasks in producing a sequence of speech acts consists in identifying entities which are to be talked about, e.g., persons, concrete objects or abstract objects, and referring to them again in the context of subsequent acts. As I already mentioned in chapter 2, one generally uses expressions representing the phrase category noun phrase (NP) for this purpose, viz. definite and indefinite noun phrases, proper names, pronouns, and "zero anaphora". For instance, if one makes a claim concerning the scientific field of genomics and goes on to make statements substantiating this claim, one will first have to introduce genomics and will then have to refer to genomics again a number of times in the course of these substantiating statements:

(50) (a) *Genomics* is a synthesis of many disparate fields, including biology, public health, engineering, computer science, and mathematics. (b) What makes *genomics* even more distinctive is that the social sciences and humanities are an integral part of the genomic revolution. (c) Philosophers, ethicists, and historians are helping to lay the foundation of the genomic revolution by pushing for and playing a role in the creation of policies and laws that will guide the integration of *genomics* into scientific practice and health care. […] (d) However, because *genomics* is an evolving science that encompasses so many different disciplines, it is hard to find one person who embodies the entire field (DeSalle/Yudell 2020, XV).

In this case, the relevant entity talked about in this paragraph – its topic –, the science of genomics, is introduced in (50a) and regularly referred to in the following substantiating statements by its proper name. At the same time a new entity related to this field of science is introduced and referred to again by the definite NP *the genomic revolution*.

[23] The word-play "which hunt" goes back to a meeting of the Chicago Linguistic Society in 1972 ("Relative Clause Festival**"**): "The Chicago Which Hunt".

Looking at the use of referring expressions from the point of view of their use in a given sequence of acts, one comes to the conclusion that it is the sequencing of acts, including the development of a topic, that is the fundamental category here and that the chain of continuous reference is just a particular feature of this sequencing pattern, which serves to implement it. Or to put it more pointedly: Coreference is an epiphenomenon of act sequencing and topic management.[24] This view is not shared by all scholars. Some authors attribute a more fundamental role to coreference and referential coherence, for instance Givón, who formulates this view as follows: "Of the multiple thematic strands of coherence […] referential coherence is the most concrete, frequent in text, grammatically conspicuous and easy to measure" (Givón 2020, 107).[25] Despite this difference in theoretical outlook most researchers working on discourse coherence will agree that coreference (or referential continuity) is an important and interesting aspect of coherence.

With the following example I will try to show how the typical act sequences and the topic management of a given text explain the use of certain referring expressions in this text. In some text types we frequently find a whole series of anaphoric pronouns signalling that the writer continues to write about a given entity, which is sometimes also the topic of the text. This is true, for instance, of obituaries. The following is an extract from an obituary for Sean Connery, in which I marked the anaphoric pronouns in these narrative passages in italics:[26]

(51) Thomas Sean Connery was born in the Fountainbridge area of Edinburgh on 25 August 1930, the son of a Catholic factory worker and a Protestant domestic cleaner. […]. *He* left school at 13 with no qualifications and delivered milk, polished coffins and laid bricks, before joining the Royal Navy. Three years later, *he* was invalided out of the service with stomach ulcers. His arms by now had tattoos which proclaimed his passions: "Scotland forever" and "Mum & Dad".

In Edinburgh, *he* gained a reputation as "hard man" when six gang members tried to steal from his coat. When *he* stopped them, *he* was followed. Connery launched a one-man assault which the future Bond won hands down.

He scraped a living any way he could. *He* drove trucks, worked as a lifeguard and posed as a model at the Edinburgh College of Art. *He* spent his spare time bodybuilding. [...]

---

[24] For this type of view see Hobbs (1979), Kehler et al (2008), Fritz (2017, 248ff.), Kehler (2019).

[25] Consequently, Givón apportions the lion's share of his treatment of discourse coherence to matters of referential coherence. A similar assumption also forms the basis of Centering theory (Grosz/Joshi/Weinstein 1995, Poesio et al. 2004).

[26] https://www.bbc.com/news/entertainment-arts-13087132. (31.10.2020).

As is typical of obituaries, the information provided is predominantly about the deceased, Sean Connery, in other words, he forms the topic of the obituary. The act structure, which is also part of the "programme" of obituaries, consists in an additive series of narrative utterances presenting events in the life of the deceased. Familiarity with this characteristic continuity of topic and the series of narrative acts "licenses" the long chain of anaphoric pronouns in the subject (or "sentence topic") position, so that the reader has no problem identifying the person to whom the anaphoric pronouns refer.

Additionally, three other types of noun phrase are notable in this stretch of text in similar functions: noun phrases with possessive pronouns (*his arms*, *his passions*, *his spare time*), which are also used anaphorically, the proper name *Connery* in the second paragraph, and the definite NP *the future Bond* (in the same sentence). The latter noun phrase, introduced for variety, also refers to Sean Connery, a reference readers will easily pick up if they know that Connery later acted in a number of James Bond films. What we have here is a good example of what Strawson called "identifying knowledge".[27]

A problem that has been extensively discussed in the literature is the question of how readers *resolve anaphors* in discourse.[28] In the above example the resolution of the anaphorical pronouns can be performed on the basis of the discourse topic (Sean Connery) and, related to this, the repeated mention of this person. In other cases, finding the antecedent for an anaphoric pronoun has to take into consideration different factors, mainly related to the accessibility and salience of the antecedent. Typical hearer/reader strategies of anaphor resolution include:

(i) looking for an antecedent agreeing in number and gender,
(ii) looking for a recently mentioned entity,
(iii) checking the grammatical role of the antecedent, preferably the subject,
(iv) looking for an antecedent that secures coherence of the sequence.

The following examples show the respective strategies:

(52) Jane and her husband travelled to Italy. *She* had never been there before.

Here the agreement in number and gender alone makes it clear that *she* refers to Jane. A somewhat more complex example is the following:

---

[27] Strawson (1964/1971, 77).

[28] See, for example, Grosz et al. 1995, Stede 2012, ch. 3, Fritz 2017, ch. 3.4, and the authors mentioned in fn. 25; see also section 7.4. For the field of psycholinguistics, Kintsch (1998, 148) characterised the situation as follows: "The literature is rich but confusing".

(53) (a) The common representation of Galileo's identity and science as alien to court values does not stem only from some author's belief in the sharp distinction between "science" and "society". (b) *It* is also maintained by those who are willing to link the development of modern science to social change but do not see the court as embodying the "good" forces of modernity (Biagioli, "Galileo Courtier", p. 1; emphasis added).

Looking in (53a) for the antecedent of the pronoun *it* in (53b), the reader will find a number of NPs that are potential antecedents, because they agree in number and gender and follow the principle of recency, namely *some author's belief*, *the sharp distinction*, *science*, and *society*. But somehow all these candidates do not fit. Therefore, the reader has to cross some distance and go back to the subject NP of (53a) (*The common representation of Galileo's identity and science as alien to court values*), which also represents the only (abstract) entity – apart from *belief* – of which it can be said that "it is also maintained". So, in this case, the combination of the subject strategy and the coherence strategy is successful.

Although the recency strategy for the accessibility of antecedents for anaphoric pronouns is statistically prevalent, as numbers presented in Givón (2005, 136) for "anaphoric distance" show, there are obviously sequential structures where long-distance pronominalisation is possible.[29]

Basically, there are two textual features that "license" long-distance pronominalisation, namely topic continuity and the continuity of a sequencing pattern. Both structural features can also be used in combination. The following extract from Stephen Nadler's book on Spinoza is an example of this technique (slightly abbreviated):

(54) (a) *Spinoza* also maintains that that if the Jewish people are "elected" in any meaningful sense, it is only a matter of their having been granted a "temporal physical happiness" and autonomous government. […] The Jews are neither a morally superior nation nor a people surpassing all others in wisdom.

(b) *We conclude therefore* (*inasmuch as God is to all men equally gracious*, *and the Hebrews were only chosen by Him in respect of their social organization and government*), *that the individual Jew, taken apart from his social organization and government, possesses no gift of God above other men, and that there was no difference between Jew and Gentile*. [...].

(c) *He* adds that if the "foundations" of the Jewish religion have not "emasculated" the minds [of the Jews] too much, they may someday "raise up their empire again" (Nadler 1999, 132; emphasis on *Spinoza* and *He* added).

[29] See the observations on this phenomenon by Barbara A. Fox (Fox 1987, 102ff.).

In (54a) Nadler presents Spinoza's view on the special status of the Jewish people, using the proper Name *Spinoza* to refer to the author. After the long-ish quotation (54b), Nadler continues his presentation of Spinoza's view by adding another element of this view (54c), tying back to Spinoza by using the anaphoric pronoun *he.* The reader has no problem identifying the person to which *he* refers, although the original reference as antecedent is textually quite distant. But as it is quite obvious that there is a continuity of both the topic (Spinoza's view) and the functional nature of the passage (the presentation of this view), long-distance pronominalisation is acceptable.

A similar technique can be found in stream-of-consciousness narratives, where pronominal reference can bridge fairly long passages of the presentation of the protagonist's mind processes. The following is an example from Virginia Woolf's "Mrs Dalloway":

(55) [Peter Walsh, who returned to London after having lived in India for some time, is walking through Regent's Park and watches a young couple obviously quarreling.] (a) And that is being young, *Peter Walsh thought* as he passed them. To be having an awful scene – the poor girl looked absolutely desperate – in the middle of the morning. (b) But what was it about? *He wondered*; what had the young man in the overcoat been saying to her to make her look like that; what awful fix had they got themselves into, both to look so desperate on a fine summer morning? The amusing thing about coming back to England, after five years, was the way it made, anyhow the first days, things stand out as if one had never seen them before; lovers squabbling under a tree; the domestic family life of the parks. (c) Never had *he* seen London so enchanting – the softness of the distances; the richness; the greenness; the civilization, after India, *he thought*, strolling through the grass (Virginia Woolf 1964, 79).

In the first passage (55a), the protagonist is (re)introduced by using his name. At the same time, it is made clear that what follows is what is going on in his mind (*Peter Walsh thought*). After two sentences the interior monologue is continued (55b) (*he wondered*), representing his train of thought about the young couple and his return to England. The final sentence of this passage (55c) again refers to him (twice) by the pronoun *he*, indicating once more that the reader is following his thoughts (*he thought*). The long gap between the beginning of (55b) and (55c) is functionally bridged by the continuity of the inner monologue, which facilitates using the pronoun after this gap. In this way it is again the functional continuity that both licenses and is indicated by the long-distance use of the anaphoric pronoun.

Seen from the perspective of the speaker/writer, a good strategy of coreference will mirror exactly the usual resolution strategies in providing good clues for the hearer/reader allowing them to find the correct antecedent without having to guess too much and to use complex inferences. Generally

speaking, the recency strategy is a good strategy, as it is in harmony with local sequencing patterns of adjacent utterances and therefore requires few additional inferences. But long distance pronominalisation is possible in cases of functional and topical continuity, a continuity that the use of the pronoun also helps to indicate. In such cases, it is the superior strategy of looking for an antecedent securing coherence of the sequence that accounts for the successful resolution of the pronoun.

I shall now continue by adding a few remarks concerning the use of definite and indefinite noun phrases in discourses.

Definite noun phrases – or definite descriptions, as they are called in the theory of reference – are often used to identify entities which are assumed to be known to the hearer/reader, either because they were mentioned earlier in the respective discourse or because they are known as part of common knowledge in the relevant community, for instance entities like the present Prime Minister, the Corona virus, and the theory of reference.

However, definite noun phrases are also used to introduce new entities into a discourse, for instance in scholarly writing. The following example is taken from the beginning of the first chapter of a book on hedging (Hyland 1998, 1):

(56) (a) *The notion of hedging* has been in the linguistic vocabulary since the term was introduced by Lakoff (1972) to describe "words whose job it is to make things more or less fuzzy". (b) It has subsequently been applied to the linguistic devices used to qualify a speaker's confidence in the truth of a proposition […].

For a reader who starts reading the book from the beginning, the NP opening the sentence (56a) (possibly) introduces a new abstract entity, namely the notion of hedging, which is the topic of the chapter and the whole book. Reference to this entity can then be taken up in (56b) by means of the anaphoric pronoun *it*.

In some cases, definite noun phrases are also used to take up reference to an entity that was introduced in an earlier paragraph and referred to there by the use of anaphoric pronouns, and which is now reintroduced in a new paragraph. In these cases, familiarity of the respective entity is presupposed but a certain distance to the original introduction is signalled.

In languages that feature *indefinite NPs*, e.g. English and German, these are frequently used to *introduce* new entities into the discourse, for instance persons in narratives:

(57) (a) Once upon a time there lived in a certain village *a pretty little country girl*. Her mother was very fond of *her*, and her grandmother loved *her* still more.

This use as a means of introduction of discourse entities is arguably based on a general use of indefinite NPs as "kind-identifying", which we also find in sentences like *What he saw was a gannet*.[30] What is remarkable about this use is that reference to the newly introduced entity can be picked up by an anaphoric pronoun, although, strictly speaking, no identifying reference was performed by the indefinite NP – the entity was simply *introduced*.

Apart from this particular use there are also other types of examples that further show the flexible use of indefinites to serve as antecedents for anaphora, of which I shall present three here, namely the use in so-called "opaque contexts", the use in certain "modal profiles", and the generalising use of indefinites in hypothetical talk.

The first type concerns opaque contexts, as they were called by Quine, for instance sentences with *I believe*. An example that is based on one of Quine's examples is the following:[31]

(58) (a) I believe he saw *a letter of mine*. (b) He should have given *it* back to me.

There are two readings of this sequence which could be paraphrased as follows:

(59) There is a certain letter of mine which I believe he saw. (b) He should have given *it* back to me.

(60) I believe he saw some letter of mine – I don't know which. (b) He should have given *it* back to me.

In the first reading a specific letter is introduced by the use of the indefinite NP and the reference is then picked up with the use of the anaphoric pronoun *it*, like in the fairy tale sequence (57). In the second reading some letter (out of a body of letters) is indicated, which is however not otherwise specified. This letter, whichever one it is, is then picked out again with *it* in (60b). The latter reading shows the remarkable flexibility of anaphoric pronouns in continuing after non-referential uses of indefinites.

The second type of example concerns the use of indefinites and pronouns in sentences featuring a modal expression:[32]

(61) (a) I ought to send my brother *a birthday present*. (b) *It* should reach him by Friday. (c) I must send *it* by express mail.

As opposed to (57a), the indefinite NP does not introduce a specific entity in (61a) but only indicates the kind of thing to be sent, namely a birthday pre-

[30] For the term "kind-identifying", see Strawson (1974, 106).
[31] See Quine (1960, 140).
[32] (61) is a variant of an example given by Karttunen (Karttunen 1976, 374).

sent. Nevertheless, an anaphoric pronoun can be used in both (61b) and (61c) to continue talking non-specifically about an item of this kind of thing. In both cases we could paraphrase the anaphoric use of *it* by *it*, *whatever present it is*. Hence, what we have here is again not coreference in the strict sense but *continuity* of non-specific indication. This continuity is based on the continuity of the modal profile, i.e., the continuing use of modals like *ought*, *should*, and *must*. That this is so can be shown by (62):

(62) (a) I ought to send my brother *a birthday present*. (b) ?I bought *it* yesterday.

This sequence is not strictly incoherent but it is pragmatically somewhat odd.[33] If one wanted to identify a definite object in (62a) and (62b) one would have to introduce it with *his birthday present* or the like, presupposing familiarity and definiteness.

The last example in this series concerns the use of indefinites in hypothetical generalisations like the following regulation:[34]

(63) (a) If a student fails *an exam*, (b) they can retake *it* at the beginning of the following term. (c) *It* will then be organised by the tutor of the original course.

In this case a hypothetical situation is presented with a possible option of action. When talking in the hypothetical mode, there is no definite reference in this case at all. What the use of the pronoun really suggests is that if we pick out one individual exam – any one – out of the choice of courses available, the rules of the regulation will apply to this exam. Hence the pronoun signals continuity, not coreference in this case as well.[35] The role of the pronoun *it* in the consequent of the conditional (63b) and the statement (63c) can be spelled out by using the following paraphrase:

(64) (a) Suppose a student fails any one of these examinations, (b) let's say phonology. (c) In this case they can retake this exam at the beginning. …

In cases like (63), the use of the anaphoric pronouns is similar to the use of bound variables in logic.

Summarising the upshot of these three examples, we can say that in all three cases the continuity indicated by anaphoric pronouns with indefinites as antecedents is based on the continuity of a type of context that is created by a

---

[33] One could, however, imagine a context in which the speaker reflects about his duties as a brother, independent of the fact that he has already bought a birthday present.

[34] This example is related to the donkey sentences invented by Geach (Geach 1968, 128f.). For a discussion of donkey sentences, see Heim (1982, ch. 2).

[35] The same is true of *a student* and *they* (used as "singular *they*").

belief report, a modal profile, or a generalising hypothetical. In other words, it is the type of sequence that licenses the flexible use of the anaphoric pronoun.

4.10 Forms of global sequencing in discourse

4.10.1 Local and global sequencing

In this chapter we have so far mainly dealt with forms of local sequencing of speech acts. There remains now the question how *global* sequencing of discourse is organised. Basically, there are three organising principles which interact in the construction of larger portions of discourse, namely:

(i) sequencing of speech act sequences (e.g. "minimal texts"),
(ii) sequencing as part of topic management,
(iii) sequencing as part of knowledge management.

As sequencing as part of knowledge management was treated in chapter 2, and sequencing as part of topic management will be dealt with in chapter 5, there is still the topic of sequencing of "minimal texts" and other sequences of speech acts missing, to which the following sections will be devoted.

4.10.2 Sequencing of big speech acts

As the basic unit for global sequencing, I shall assume *paragraphs* (of differing size). A unit of this size can be, for instance, used to realise a description, which consists of a sequence of statements like the following description of the 19th century Fulton market (D'Costa 2011, 64):

(65) (a) In the early 1800s, a large open air market was set up on Fulton street. (b) The market sold the staples of everyday life, including coffee, shoes, stationary, books, ice cream, and pistols, as well as fruits, vegetables, and dairy products. (c) Located near the East River, its proximity to the water also meant plenty of fresh fish. (d) But the fishmongers were only a section of the market at this time. (e) The market's location also meant plenty of customers: the Fulton Ferry could steam residents over from Brooklyn to add to the trade. […] (f) The market was also a place to get a meal. There were several push-cart vendors and oyster stalls that served lunchtime crowds. (g) Market days were a regular feature of social life.

Such a paragraph is not only a sequence of speech acts (a) to (g) but could itself be considered a big speech act, namely a description. In a similar fash-

ion, bits of narrative or argumentation of the type presented in the section on "minimal text types" could be described as big speech acts in the sense in which I want to use the expression. The same is true of explanations, justifications, accusations, directives, suggestions, and evaluations, which come both in small versions, performed by using one single sentence, and in big versions, performed as whole sequences of speech acts. Types of speech act performed by utterances of paragraph size which occur mainly in the form of "big" speech acts include, for instance, summaries, abstracts, problem descriptions, and so-called advance organisers.

The next point to clarify is the question in what way paragraphs and the big speech acts performed by them can be combined. As an example, I give a narrative passage that could follow directly after the description (65):[36]

(66) (a) In 1821, a fire swept through the neighborhood. (b) Because the market was a vital part of the neighborhood, (c) it was rebuilt by the city in 1822 (d) and the fishmongers were assigned to a side of the new building. (e) However, other vendors – particularly the butchers – began to complain about the fish sellers, claiming that the smell was bad for business. (f) They petitioned to have the fishmongers moved, and were successful: (g) The fishmongers were moved to a shed, but they took a good portion of business with them. […]

This combination of a descriptive passage with a narrative passage – in this order – is so frequent that one could feel justified in assuming a *sequencing pattern for big speech acts*, namely "description & narrative". It is this kind of sequencing that fundamentally contributes to the global functional organisation of texts. As for the internal structure of the narrative, it is worth noting that among the event-presenting sentences there is also the explanation (66b), which is also not an unusual contribution to a narrative passage.

Other well-established sequences of big speech acts are "narrative & explanation", "narrative & justification", "description & summary", "problem description & description of a solution", "presenting a thesis & presenting evidence for the thesis" etc. "narrative & explanation" is a frequent pattern in historical research, e.g. in many texts of historical semantics, where first the historical development of a meaning change is presented and then an explanation of the change is attempted.[37] "narrative & justification" occurs, for instance, in cases where somebody tells a story of his problematic activities and then presents a battery of reasons why he could not have acted otherwise.

Sequences of this type can be further combined with each other, and with

[36] After the above description, the original text by D'Costa inserts a quotation giving some extra information on the social life of the market.

[37] On the forms of representation used in historical semantics, see Fritz (2020a).

other patterns, which allows the production of even more complex texts like the ones following the pattern "description & narrative & explanation & evaluation". Generally speaking, it is this system of textual elements and their constellations that fundamentally contributes to the global coherence of many functionally complex texts.

As for the sequencing order in these patterns, this is frequently determined by principles of knowledge management. In order to be able to tell a story it is sometimes necessary to provide relevant knowledge in form of a descriptive textual element. The same applies to the explanation of an event: Before explaining the event, one has to tell the story – if it is not already common knowledge. There are, however, also presentation strategies that consist in intermingling elements of the different patterns, e.g., the strategy of providing bits of relevant knowledge in a narrative "as you go along".

In addition to these general elements and principles of global sequencing there are also very specific routines of the combination of speech act patterns with topic structures and principles of knowledge management that have developed for certain purposes in certain communities. Such global patterns and routines are called *text types* or *genres*, to which I now turn.

4.10.3 Text types and global sequencing

As I already mentioned in chapter 2, text types or genres, as some people prefer to say, are basic structures of whole texts, which include the functional structure, the topic structure, the knowledge building structure, communication principles, and the utterance form of texts. As such they provide criteria for the production and assessment of the global coherence of texts. Examples of text types with quite a variety of textual features are: news articles, crime fiction, scholarly reviews, medical package leaflets, cooking recipes, penalty charge notices, and weather forecasts.[38]

Depending on the point of view from which one looks at the concept of text type one can describe text types as text patterns, text routines, or text models. Considering them as text patterns, one will emphasise their structural properties, e.g., their typical functional and thematic elements. Describing them as text routines, one emphasises the fact that they are established elements of a communicative practice, which have developed historically as solutions to communicative tasks and which are frequently followed unre-

[38] The topic organisation of two text types (medical package leaflets and research papers) will be discussed in section 5.3.3.

flectingly by members of a given community.[39] Considering them as models or templates, they are seen as standards which are to be followed, which are sometimes codified in guidelines, and which have to be learnt by novices to the respective communicative practice. Generally speaking, text types range from vague ideas and outlines of how to write a specific kind of text to standardised and strictly normative rules for the production of certain texts. And it is worth keeping in mind that some texts are modelled on a *combination* of text types, e.g., the combination of historical narrative and sociological analysis.[40]

For the purpose of presenting some general features of text types I choose the *scholarly book review* as an example, a type of text which is frequently published in scientific journals, mostly in the humanities but also in some of the sciences.[41] Looking at the functions performed by various review-like texts and the constellations of functional elements to be found in individual reviews and review-like texts, we find that typical functions like summarising a piece of scientific work, critically discussing such a work and evaluating it are spread over a variety of text types and media formats, e.g. reviews of literature that form part of PhD-theses, pre-publication peer-reviews, comments in Open Peer Review, review-like texts published on mailing lists or as blog posts.[42]

Book reviews in the narrower sense of the word show (and showed historically) a wide array of functional and thematic elements in addition to the ones mentioned before, e.g.:[43]

---

[39] For the analysis of genres as "solutions to communicative problems", see Luckmann (1986); Knoblauch/Luckmann (2004).

[40] Concerning individual text types, there are several corpus-based studies performed within an action-theoretical framework, e.g., Schröder (2003) (text types in newspapers), Wolańska-Köller (2010) (cooking recipes), Fritz (2011) (posts on science blogs), Cheng/Gloning (2017) (verbal portraits in newspapers), Kaltwasser (2019) (contributions to online forums), Bucher (2019) (Twitter microblogs), Bucher/Boy/Christ (2022) (YouTube science videos).

[41] For a more detailed analysis of the practice of scholarly reviewing, see Fritz (2020b).

[42] In addition to review-like texts, science blogs frequently present quite a variety of (small) text types, from descriptive to polemical texts (see Fritz (2011).

[43] For an analysis of the basic functional elements ("moves") in English scholarly reviews (in chemistry, linguistics, and economics), see Motta-Roth (1998). For the combination of speech act types, thematic structures, and knowledge management in science texts, see Gloning (2020); for theatre reviews, see Gloning (2008).

(i)	an introduction of the topic of the book under review,
(ii)	an outline of the organisation of the book (topics and/or argumentation),
(iii)	a description of the relevant state of the art,
(iv)	a summary of relevant discussions in the respective community,
(v)	remarks on the status and qualification of the author,
(vi)	a presentation of competing theories,
(vii)	a comparison with work on a related topic,
(viii)	an evaluation of individual aspects of the book
(ix)	an attempt at a refutation of the author's theses,
(x)	a polemical attack on the author (or the school he represents),
(xi)	a presentation of extra data collected by the reviewer,
(xii)	an attempt to adjudicate in a controversy of which the work under review is a part,
(xiii)	a general evaluation of the book,
(xiv)	a statement of remaining research desiderata.

It is true that there has been a process of standardisation in various scientific disciplines restricting the accepted range of functional elements in reviews, a standardisation which evolved both "naturally" by the selection of useful textual strategies in the course of the professionalisation of these disciplines and "normatively" by the influence of explicit review guidelines issued by the editors of journals or even by the introduction of "structured book reviews", which are prevalent in some medical journals. However, in a sufficiently large corpus of reviews we find all these moves and others. And this latitude also makes sense functionally, as all these types of elements are potentially useful contributions to scientific information and exchange. Thus, the text type of scholarly book review really stands for a family of text types related by various degrees of family likeness. As this is generally the case with text types, it is always an interesting point to explore the degree of standardisation or room for manoeuvre characteristic of a particular text type.

As for the *sequencing of the basic acts*, it has been noted that a general evaluation of the book under review is standardly used as a closing move, but sometimes also as an opening move.[44] The evaluation of specific aspects of the book, however, is often closely linked to the description of these aspects, so evaluative elements tend to be dispersed all over the text.

Concerning the *topic structure* of reviews, one often finds a movement

[44] For the role of evaluative acts in scholarly reviews, see Hyland (2004, ch. 3) and contributions in Hyland & Diani (2009).

from more general to more specific aspects of the book, a sequencing strategy that contributes to a systematic structuring of information. This is, for instance, true of the description of the general topic, followed by the subtopics of individual chapters.

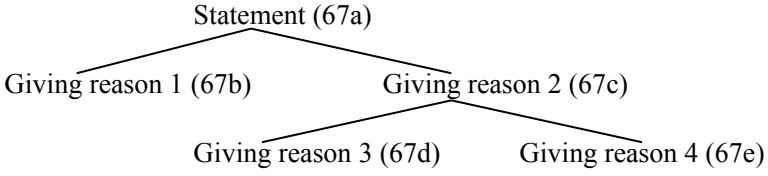
For many text types, there exist *communication principles* that are characteristic of this text type. In the case of scholarly reviews, some of these principles tend to be mentioned in guidelines for reviewers, e.g., the principle of comprehensibility, the principle of informativity, the principle of brevity, and principles of politeness, especially face-saving principles. In the guidelines for a history of science journal, we find, for instance, the following advice: "reviewers should avoid inflammatory language and aim for criticism that is fair, balanced, respectful, and focused on the work written, not on the author."[45] In spite of the general acceptance of such principles, reviews sometimes spark off controversies which develop by responses to reviews and replies to responses.[46]

4.11 Functional hierarchies and patchwork texts

It has frequently been observed that in many texts we find hierarchical functional structures like the following from an argumentation:[47]

(67) (a) I can't help you paint your room next week. (b) I am not feeling too well. (c) And I have no time next week. (d) I have to study for an exam. (e) And my parents will be visiting me.

The structure of this short argumentation can be described as follows: A makes a refusal by stating (67a) and then supports this statement by giving the reasons (67b) and (67c). In addition, A supports (67c) by giving the reasons (67d) and (67e). Schematically the structure can be presented as follows:



[45] https://www.ircps.org/aestimatioguidelines-reviewers (04.04.2021).

[46] For an analysis of an Early Modern example, see Fritz/Glüer (2018).

[47] Based on an (invented) example from van Eemeren/Grootendorst/Snoeck Henkemans (2002, 71f.). The authors also present a chart showing the structure of a "combination of various types of complex argumentation".

The rightmost branch shows a hierarchy with three levels, structured by the sequencing pattern "(67a) is supported by (67c) & (67c) is supported by (67e)". This kind of functional structure is quite frequent with all kinds of giving reasons and argumentations, e.g., explanations, justifications, refutations, giving evidence for empirical statements and giving examples for generalisations etc.[48] As such, these structures are frequently embedded in larger structures, which may or may not be hierarchically structured, as, for instance, longer passages in narratives, news reports and descriptions of historical situations.

In some cases, especially in short texts with a single purpose, for instance requests or formal injunctions, such a hierarchical structure may determine the whole text. This observation lead some researchers to the assumption that texts are *generally* built according to this hierarchical functional model (e.g. Brandt/Rosengren 1992, Mann/Thompson 1988) and that this kind of structure determines the *whole* text.[49] This fairly strong assumption is, in the case of Rhetorical Structure Theory, mirrored in the claim "that the conditions of completeness, connectedness and uniqueness taken together are sufficient to cause RST analyses to be trees" (Mann/Thompson 1988, 249).[50] This view of the functional structure of texts also determines their concept of coherence. A text is coherent in this sense if the embedded structure of coherence relations covers the whole text.

This conception of the functional structure of texts is, however, mainly true of small, "well-behaved" texts or parts of larger texts, whereas larger, more complex texts tend to show a more complicated structure which is mainly characterised by the interplay of functional sequences and topic structures. From the point of view of the practice of text annotation, Stede perceptively noted: "… experience with annotation experiments points to the recommendation of confining coherence-relational structure to relatively small units of text, roughly paragraphs, rather than forcing annotators to glue large trees together on a level of description that might be more adequately described by content zone or topic structure" (Stede 2012, 126).[51]

The following are some additional arguments against the assumption of a

---

48 See the minimal forms of argumentation introduced above.

49 Swales' "moves and steps" model assumes at least partially hierarchical functional structures (Swales 1990).

50 The authors do, however, later concede that "certain text types characteristically do not have RST analyses. These include laws, contracts, reports 'for the record' and various kinds of language-as-art, including some poetry" (Mann/Thompson 1988, 259).

51 Stede's "content zones" are equivalent to Swales' "moves" (see Stede 2012, 13).

consistently hierarchical functional structure of texts in general. In the first place, there are text types that are characterised by "flat" functional sequences like, for instance, certain types of news reports, called "additive report" ("additiver Bericht") in Schröder's analysis (Schröder 2003, 208), which consist of a concatenation of (small) event reports referring to the same major event. Examples of text types that are dominated by a hierarchical *topic* structure, containing in general only local configurations of hierarchical functional sequencing, include descriptive texts like encyclopedia articles (e.g. Wikipedia articles) or the field guides mentioned above.[52] Such internal local configurations consist, for instance, of explanations of concepts or of arguments for a certain theoretical viewpoint.

In some text types we have a small number of functional centres embedded in a central topic, e.g., descriptions and evaluations (or: comments) in some types of review. But, as I mentioned earlier in this chapter, scholarly book reviews frequently show a more extended array of functional elements. In addition, it is worth noting, as Taboada et al. (2009, 63) remarked for the case of film reviews, that "it is often the case that both description and comment are present in the same paragraph".

Another type of text where no unified hierarchy of functional structure is to be expected is formed by the wide variety of multifunctional texts, including textual portraits of touristic destinations (see Ermakova 2015), film reviews (see Stegert 1993), essayistic texts, and hagiographies (see Fiona Fritz 2022). In such texts we find individual clusters of functional elements with (partly) hierarchical structures as "islands" embedded in various topic structures. Informally, texts with global structures of such a more fragmented format could be termed "patchwork texts" or "bricolage texts". Not surprisingly, they present a considerable challenge to the analysis.

4.12 Speech act sequencing patterns and coherence relations

Among the various approaches to the explanation of discourse coherence a number of researchers use so-called coherence relations to account for the connectivity between spans of text and the structure of texts. The most well-known approach of this kind is Rhetorical Structure Theory (RST), presented

[52] For encyclopaedia entries, van der Vliet/Redeker (2014, 35) observed a rather flat functional structure, described in terms of RST relations. For topic structures, including hierarchical structures, see the following chapter 5.

in the 1980s (e.g., Mann/Thompson 1988, Taboada/Mann 2006).[53] As an example of the use of coherence relations – called "rhetorical relations" in this approach – for the description of text sequences I quote the introduction of the RST relation JUSTIFY (Mann/Thompson 1988, 252):

> The following short text, from the electronic bulletin board at ISI, provides an example of the **Justify** relation:
>
> 1. The next music day is scheduled for July 21 (Saturday), noon-midnight.
> 2. I'll post more details later,
> 3. but this is a good time to reserve the place on your calendar.
>
> In this text, Units 2-3 are in a **Justify** relation with Unit 1. They tell readers why the writer believes he has the right to say Unit l without giving 'more details', in particular, without giving the location of the music-day event.

In their paper the authors suggested a general system of 24 rhetorical relations, from EVIDENCE to ELABORATION. Other authors working in this framework suggested systems of up to 70 relations (e.g., Hovy/Maier 1995, Marcu/Romera/Amorrortu 1999).

As part of their analytical tool-kit the authors further use a pair of concepts, *nucleus* (N) and *satellite* (S), which serve to differentiate the "central" text span (the nucleus) and the one "supporting" it (the satellite). In the above example, unit 1 is the nucleus, whereas and the units 2-3 form a satellite.

In the present section I intend to clarify the relationship between typical coherence relations of the RST type and sequencing patterns of act types. As my starting point I shall take the claim that, arguably, (at least some) coherence relations are *based on* sequencing patterns, which is equivalent to the assumption that speech act sequencing patterns are *more fundamental* than coherence relations. A further assumption, which I shall try to justify, is that for some/many purposes of analysis it is preferable to use the concept of speech act sequence instead of using coherence relations.[54] A major disadvantage of the rhetorical-relations approach, which I shall however not dis-

---

[53] Readers not specifically interested in RST and coherence relations may want to skip the next two sections and go on to chapter 5. For a critical discussion of RST, see Fritz (2014), Fritz (2017, ch. 14).

[54] The basic idea was presented in Fritz (1982, 51-55). In the context of the analysis of multimodal documents, this critical observation was also made by Bucher (Bucher 2017, 103). A similar idea, albeit in a different theoretical framework, was voiced by Asher and Lascarides, who suggest "viewing rhetorical relations as *speech act types*" (Asher/Lascarides 2003, 112). Stede and Peldszus combined a coherence relations approach with a speech acts approach by proposing "an inventory of categories of illocutionary status for labelling the spans that are being connected" (Stede/Peldszus 2012, 214).

cuss here, is that reference and coreference cannot be represented by this approach.[55] A speech-acts approach, by contrast, systematically takes into account aspects of reference (e.g., Searle 1969, ch. 4).

I shall now try to show some of the advantages of a speech-acts approach, using as an example the rhetorical relations *Justify* and *Evidence* presented in Mann/Thompson (1988, 251f.) The "relation definitions" of these rhetorical relations are as follows:

> **Evidence** and **Justify** form a subgroup; both involve the reader's attitude toward the nucleus. An **Evidence** satellite is intended to increase the reader's belief in the nuclear material; a **Justify** satellite is intended to increase the reader's readiness to accept the writer's right to present the nuclear material.
>
> **Justify**
>
> *constraints on N*: none
> *constraints on S*: none
> *constraints on the N + S combination*: R's comprehending S increases R's readiness to accept W's right to present N
> *the effect*: R's readiness to accept W's right to present S is increased
> *Locus of the effect*: N
>
> **Evidence**
>
> *constraints on N:* reader might not believe N to a degree satisfactory to writer
> *constraints on S:* reader believes S or will find it credible
> *constraints on the N+S combination:* reader's comprehending S increases reader's belief of N
> *the effect:* reader's belief of N is increased
> *locus of the effect:* N

There are several problems with these defining constraints, of which I shall briefly indicate three here:

(i) One problem concerns the "constraints on S", i.e., the conditions to be fulfilled by the writer (W). "reader believes S or will find it credible" is an effect analogous to what speech act theorists have called a perlocutionary effect. However, this effect cannot be a condition for the characterisation of the text span (N) as a case of *Evidence*, because the effect on the reader is not something the writer can guarantee and also nothing the annotator can know. What the writer *can* guarantee – and what the annotator can assume in his description of the span as a case of *Evidence* – is that the writer has the *intention* to increase the reader's belief in whatever is asserted with N – if we take N to be

[55] Further disadvantages are that this approach lacks concepts like topic, context change, and the *by*-relation.

used as an assertion. Analogously, instead of the defining an "effect" it should be the *intention* of the writer that is at stake here.

(ii) A second problem consists in the lack of defining "constraints" for N and S in the case of *Justify*. It is rather puzzling that there should be no information on the "*constraints*" on N and S parallel to those given for *Evidence*. Following the format used for the definition of *Evidence* one would expect for *Justify* a constraint on N like "The writer assumes that the reader might consider the writer's reasons for asserting N not satisfactory". And for the constraint on S: "The writer intends S to provide satisfactory reasons for the reader to accept/believe N".

(iii) The authors acknowledge that *Justify* and *Evidence* form a subgroup, they do however not make explicit what their relationship is. What I should like to show is that these two relations are closely related in a way that is not indicated by the definitions. One could, for example point out that one of the best reasons for the reader's "readiness to accept W's right to present N" (constraint for *Justify*) is his "believing S" (constraint for *Evidence*).

(iv) The description lacks the concept of *by*-relation, which, as I shall show, is necessary for a satisfactory description of the patterns "justifying" and "providing evidence".

In order to show the relation between *Justify* and *Evidence*, I shall have to insert a few reflections on the status of such constraints. What these defining constraints seem to provide is a number of conditions for the performance of the relevant speech acts. In the case of "providing evidence for p" a commitment of the writer could be the assumption that "the reader might not believe p to a degree satisfactory to the writer", which corresponds to the „constraint on N" for *Evidence*. Thus, the formulation of these rhetorical relations could be read as a kind of shorthand for the description of the respective speech acts.

Now if we look at the speech acts types of "justifying" and "providing evidence" we find that they are fairly closely related. Let us start by looking at "justifying". "Justifying" is a communicative act type that is mainly related to acts and beliefs.[56] One fundamental type of justifying consists in "giving reasons", so we can assume a general pattern "justifying X *by* giving a reason (or reasons) for X". As for "providing evidence", one could start with the following assumption: "Evidence, whatever else it is, is the kind of thing which can make a difference to what one is *justified* in believing [...]" (Kelly 2016, 1). Therefore, we can obviously also assume an act pattern of the type

[56] See Kasher 1987: "Justification of speech, acts and speech acts".

"justifying X *by* providing evidence for X". And we could further assume a sequencing pattern of the type "asserting that p & justifying the assertion that p *by* providing evidence for p". An example of this kind of sequence is:

(68) (a) There was somebody in the house. (b) Three witnesses heard the noise.

The evidence (68b) justifying the assertion (68a) addresses the belief that there was somebody in the house, to which the speaker/writer is committed by making the assertion (68a). But in justifying this belief the speaker also justifies the assertion of the belief. In this way, we can show a close systematic relationship of these patterns that is not covered by the relation definitions of RST.

A similar analysis could be given for the example Mann and Thompson give for their rhetorical relation *Evidence*:

(69) (a) The income tax program as published for calendar year 1980 really works. (b) In only a few minutes, I entered all the figures from my 1980 tax return and got a result which agreed with my hand calculations to the penny.

The utterance (69b) provides evidence for the satisfactory working of the programme and it also serves to justify the speaker's praise of the income tax programme.

For the practice of text annotation, the problem observed just now may lead to a systematic difficulty that is well-known from RST annotations. In cases where the relevant text span could be described according to the sequencing pattern "justifying the belief that p *by* presenting evidence for p" the RST annotator is forced to decide whether the respective text span should be annotated with *Justify* or with *Evidence*, a situation that leads to a dilemma. But this dilemma is not caused by the incompetence of the annotator; it is a result of the unsatisfactory analysis of evidence and justification and the lack of the concept of *by*-relations.

Now we still have something to elucidate about the act pattern of "justifying". As I said above, we can also distinguish a pattern of "justifying X by giving a reason (or reasons) for X". Now, in some cases the availability of reliable evidence for X can certainly be considered a good reason for accepting X, as in the examples (68) and (69). Therefore, one could, in these cases, assume an act pattern of the type "giving a reason *by* providing evidence". The upshot of these reflections is that the act pattern "justifying" is really a family of related patterns, the members of which might be worth distinguishing or not, depending on the aims of the enquiry at hand.

4.13 How many speech act sequencing patterns are there?

After this excursus on coherence relations, I will now briefly turn to a thorny question of the theory of coherence, namely how many types of act sequences there are. This type of question mainly comes up in the context of methods for an economical and reliable analysis and annotation of texts. A much-discussed variant of this problem is the question of how many coherence relations should be assumed, so I shall start my discussion obliquely from the latter question.

A well-known answer was given by Mann and Thompson in their much-cited paper on Rhetorical Structure Theory (RST, Mann/Thompson 1988). As I mentioned in the last section, they suggested a general system of 24 rhetorical relations, from EVIDENCE to ELABORATION, whereas other authors in this framework suggested systems of 54 relations (Marcu 2000) or 70 relations (Hovy/Maier 1995, Marcu/Romera/Amorrortu 1999). On the other extreme, Asher and Lascarides felt that for their (truth conditional) purposes seven discourse relations would be enough (Elaboration, Explanation, Result, Background, Contrast, and Parallel; Asher/Lascarides 2003, 145). Thus, there appeared to be a problem of standardisation of the canon of rhetorical relations, as Alistair Knott noted: „Apart from anything else, there is a danger that unless a standard set begins to emerge, the research programme based on coherence relations might grind to a halt" (Knott 1996, 37). This problem arises in particular if one wants to compare analyses of different types of discourse by using the same system of annotation.

In the standard treatments of RST this problem is not given much attention. There were, however, several authors who specially addressed the question (e.g. Hovy/Meier 1995). Starting from the "parsimonious" position of Grosz and Sidner, who rightly claimed that trying to identify the "correct" set of discourse relations is a doomed enterprise because there is no closed set (Hovy/Meier 1995, 4; cf. Grosz/Sidner 1986, 179) and therefore restricted themselves to two global categories (dominance and satisfaction-precedence), Hovy and Meier reached the conclusion that "the solution [to the problem] we propose is to use just as many relations as are required for the task being done" (Hovy/Meier 1995, 5). This is in accordance with the general principle that the granularity of a description should be determined by the purpose of the respective description.[57] And it also mirrors the fact that

[57] In the practice of annotation there is of course a trade-off between the specificity of description to be achieved by using a large inventory and the simplicity of annotation with a small inventory of categories.

researchers analysing specific types of text have tended to introduce coherence relations of their own, relevant to their particular text type. This is, for example, true of the paper by Rösner and Stede, who worked on "customizing RST for the automatic production of technical manuals", showing that the original set of RST relations is not specific enough for practical *generation* purposes and suggesting a number of new relations they found useful in this particular domain, e.g. the relations of "preparation" (for an action) and "sequences of steps" for the description of steps the user of a technical device is instructed to perform (Rösner/Stede 1992).[58]

The upshot of these observations is that we have a dilemma created by a pre-defined uniform inventory of categories that is not attuned to different purposes of analysis. If we want to carry out a uniform analysis of coherence in different types of text from different domains, which might be a useful exercise, it seems natural to use a unified system of coherence relations. But this comes at a price, which can be made explicit by the following two objections:

(i) Using, for example, the RST-inventory of rhetorical relations for an annotation and analysis of a corpus of texts from different genres, we might end up with the result that in all these texts coherence is based on discourse relations from the RST-inventory and only those, which is a clear case of circularity.

(ii) When analysing texts of a certain genre, e.g., scientific controversies, we will probably find types of argumentation that are not provided for in the given inventories of categories. So we are faced with a lack of specificness. Now if our special interest in this particular study is exactly focused on different types of argumentation, e.g., different strategies for backing claims or the use of different kinds of evidence, we will probably have to extend our set of descriptive categories and even introduce new ones that only emerge in the course of our analysis. Restricting our analytic inventory to a given set, e.g., the RST set of rhetorical relations, would force us to give up the essential aims of our study, so this is really no option.

Both these objections, circularity and lack of specificness, are rather grave and should be avoided. To solve this dilemma, a technical solution might help. If we want to be free to introduce relevant specific categories and at the same time to preserve the generality that might be attained by using a unified inventory of categories we could use a multiple stand-off annotation, which

[58] Similar relations, which are characteristic of this text type, were introduced in a more recent paper dealing with instructional texts e.g. "preparation:act" and "step1:step2" (Subba/Di Eugenio 2009, 568).

is already used for multi-level annotations anyway. This could provide the best of both worlds and, in addition, facilitate a comparison between different annotation solutions.

The result reached by these reflections on coherence relations basically also applies to the panorama of types of speech act sequences one would want to use for the analysis of different kinds of discourse. Fundamentally, one should use just as many speech act types and sequencing types as is required for the descriptive task being done, as mentioned above.

Thus, before deciding which repertoire of descriptive categories we want to use we have to answer at least the following three questions:

(i) What is the purpose of the description/analysis?
(ii) What is the scope of the description/analysis (e.g., concerning types of discourse)?
(iii) Which degree of granularity of description is necessary for this particular purpose?

This repertoire will, in any case, be richer and more finely differentiated than that of general descriptions of speech act types like Vanderveken's (1990, ch. 6), who gives an inventory of ca. 70 types. This is due not only to the fact that for many speech act types there are variants like those mentioned above in the case of "justifying" – e.g. different types of argument – but also to the fact that we find variants like different degrees of commitment and other kinds of specification of statements, suggestions or evaluations – e.g. hedged statements, cautious suggestions or evaluations – and, of course, many differences of propositional content that may be relevant for the analysis of particular discourses. In addition, there are, as mentioned before, many types of speech act that are related to the organisation of discourse – e.g., "announcing a topic" and "concluding an argument" – or to other aspects of a discourse (e.g., "indicating the state of common knowledge"). All these speech acts can also appear in sequences with other speech acts, e.g., explanations. Furthermore, there are many sequencing patterns that are characteristic of individual text types, e.g., news articles or technical manuals. And finally, we must not forget that speech act sequences have a historical dimension: in earlier historical periods we find different patterns, for instance, in the case of explanations and justifications, and we can see that innovative sequences are permanently being created (e.g., multimodal sequences), some of which may later become routinised and established as sequence types.

So, to return to the starting point of these observations, the question of how many speech act sequence types there are has no definite answer, at least no simple answer. One could, of course, try to make an inventory of the

speech act sequence categories actually used in many different analyses, but that is more than I can do at this point.[59]

4.14 Corpus studies on sequencing patterns

Surveying the state of the art in studies on discourse coherence, we find that only a restricted sector of resources used in creating coherence has been regularly analysed using *corpus methods*. Two such resources stand out, namely the use of referential means for "entity coherence" (forms of coreference, pronominalisation, "anaphoric distance"; e.g., Poesio et al. 2004, Givón 2020, ch. 6) and the use of connectives to indicate coherence relations (e.g., Marcu 2000, Taboada 2006, Prasad et al. 2008, Stede 2012).[60] This aspect of the field also accounts for the fact that *quantitative information* on discourse structures is mainly available for these factors.

As for coherence relations, we basically find two approaches to corpus analysis, namely manual annotation of a corpus and automatic detection of coherence relations. Detecting and annotating coherence relations (and, mutatis mutandis, sequencing patterns) is simplest when they are explicitly indicated on the text surface by connectives (and other lexico-grammatical means). It is therefore not surprising that corpus studies of discourse structures should have concentrated on these types of relations, (e.g., Breindl/Walter 2009, Prasad et al. 2010). *Implicit relations*, which are generally inferred by hearer/readers on the basis of their contextual and commonsense knowledge and their knowledge of sequencing patterns, present quite a challenge to the analysis and have mainly been dealt with by manual annotation. The problem can be illustrated by the following example given by Peldszus and Stede in a paper on argumentation:

(70) The book never appeared. The publisher had gone bankrupt.

As the authors noted, "… the causal relationship can only be inferred by the reader on the basis of his knowledge on the origin of books and the consequences of insolvencies" (Peldszus/Stede 2013, 22). This means that a relia-

[59] A heuristic for charting the "geography" of speech act sequence types is given in Fritz (2017, 177ff.).

[60] In addition, there is corpus-based research on automatic topic detection and the topic-subtopic structure of texts (see section 5.1); in the section on text types (section 4.10.3) I also mentioned a number of corpus-based studies analysing global coherence structures within an action-theoretical framework (e.g. Schröder 2003).

ble automatic analysis of implicit coherence relations is difficult to achieve and is still not feasible today.[61] As a typical problem for his surface-based algorithm, Marcu mentioned the following: "The rhetorical parser is unable to recognize "purely" intentional relations, such as EVIDENCE, which are seldomly marked" (Marcu 2000, 439). A similar problem has been observed in the case of *implicit evaluations*, which in some genres play an important role for discourse coherence and which can only be determined on the basis of knowledge about relevant evaluation criteria and factual knowledge about the objects to be evaluated (e.g., Benamara et al. 2017). Seen from the perspective of the present book: Some of the most intriguing aspects of coherence turn out to be the most difficult to analyse by means of computational methods, and corpus analysis is therefore costly in terms of time and human effort.

[61] For a brief survey of problems involved in detecting implicit coherence relations, particularly in the case of larger textual spans, see Stede (2012, 110ff.). Stede's book is a useful introduction to methods of discourse processing.

5. Topic structures and topic management

5.1 Topics and their function in discourse

There seems to be no doubt: one of the basic organising principles we follow in having conversations and writing texts is the orientation towards topics. Following this principle fundamentally contributes to the achievement of coherence. The concept of "topic" is an everyday category that we explicitly use when we "set" a topic, when we report on someone having talked *about* a topic or having "raised" a topic, or when we complain that someone was completely *off-topic* in a discussion. And indeed, some researchers feel that the concept of topic is the basic concept in terms of which to explain discourse coherence (cf. Giora 1985). Strangely enough, this everyday concept is not easy to define and to operationalise. This insight was stated by Givón as follows: "Thematic continuity is the overall matrix for all other continuities in the discourse. It is the hardest to specify, yet it is clearly and demonstrably there" (Givón 1983, 8). In fact, some researchers working on formal theories of discourse shun away from using this concept at all. In the words of Nicholas Asher: "some see topics as creatures of darkness – a viewpoint I can readily understand" (Asher 2008, 46).[1] As will become clear in the course of this chapter, this is not the position I take in this work.

In the following sections I shall try to clarify the concept of topic as I use it and at the same time indicate the role topic management plays for the creation of coherence.

The first step in clarifying the concept of topic usually consists in saying that the topic is something that a text or a conversation is *about*. However, as "aboutness" is itself a concept very much in need of clarification, there is not very much gained by this step. One could say that a text is about X if the centre of interest of the discourse is X.[2] One possible operationalisation of this idea is based on the assumption that the centre of interest is indicated by the fact that the expression used to formulate the topic of a discourse actually occurs regularly or at salient points in the discourse. This is indeed frequently the case, e.g., if we write an article on the Covid pandemic it is quite likely that

---

[1] Despite this attitude, he uses the concept of topic, mainly for the analysis of narratives and extended descriptions, in Asher (2004).

[2] An early discussion of the idea of topic as the "centre of interest of a statement, the idea of what a statement could be said, in this sense, to be about" can be found in Strawson (1964/1971, 92f.). Strawson's reflections on aboutness of statements partly also apply to discourses.

we should use the expression *Covid* frequently or at salient points in the article. In academic writing, for instance, it is certainly a good *writing strategy* to use potential "topic expressions" at crucial points of a paper, e.g., in the title, the abstract, the summary or at the beginning of important paragraphs.

But the frequency or even the mere presence of individual "topic expressions" cannot be a necessary criterion for topichood, as we can easily write an article about Covid using mostly expressions like *coronavirus* or *the virus* or *the pandemic* and maybe only using *Covid* once or twice and in completely inconspicuous places. In this case, the readers may still agree on saying that it is an article about Covid. Another problem is the following: In saying that a text is about X, we normally use a noun phrase to identify the topic, for instance by saying *the text is about NP* (e.g., about recent political debates), but it is possible that most sentences directly contributing to the topic of the text are characterised by the use of certain *verbs*, like *disagree*, *discuss*, *argue*, *object*, *convince* or *prove*, whereas the NP is only rarely used. Finally, a topic can be named by different expressions, e.g., *the 45th president of the United States*, *the worst president of the US ever* or *Donald Trump*, and potential topic expressions may be polysemous, in which case the occurrence of an expression like *bank* alone does not determine whether the text is about a banking institute or a river bank. In these and many other cases it is necessary to have additional information in order to be able to recognise the real topic of a text. More generally, one can say that the topic does not lie "objectively" in the surface structure of the text but is an aspect of the meaning of the text.

Together, these properties of topics – and others – present a considerable challenge to attempts to perform "automatic topic detection", an interesting project in the context of "text mining". As can be inferred from the features of topics mentioned just now, a simple word count with the aim of detecting a (possible) keyword on the basis of its frequency is bound to fail as a reliable method of topic detection.

One way out of this impasse consists in using additional aspects of the texts to be analysed. Recognising that certain writing strategies, e.g., the scholarly ones mentioned above, place topic-indicating expressions in typical positions of the text, some information scientists tried to develop a method for the automatic identification of topics by "locating the likely positions of topic-bearing sentences based on genre-specific regularities of discourse structure" (Lin/Hovy 1997, 283). This method, they argued, was much more accurate than alternative methods based on simple word counting.

A second type of strategy, which is frequently used today, consists in determining whole *clusters* of words that co-occur with potential topic words in texts (or paragraphs of texts) or *sequences* of pairwise related words in a text, so-called "lexical chains". If, for instance, the expressions *bombers*, *missiles*,

explosions, *destruction*, *infra-structure*, *casualties* occur in a newspaper report, the topic of the report is likely to be an air raid. (But the article may also be implicitly about the cruelty of the aggressor.) Such a cluster then forms a kind of "topic signature" (see Lin/Hovy 2000) of the text or the paragraph. In this example one could also show that the expressions occurring in the sequence are pairwise related by co-occurrence relations and therefore form a "lexical chain". Lexical chains "serve as a record of a topic's (or subtopic's) presence in a text" (Stede 2012, 23). In these methods, a topic is either defined as a distribution of words, namely the set of expressions included in the cluster or chain or, more cautiously, as an expression that can be inferred on the basis of the cluster or chain.[3] This means that they generally suffer from being restricted to information from the text surface, excluding the additional knowledge that is frequently needed to understand the topic of a text. Therefore, they tend to produce rather coarse-grained results and cannot be used for more sophisticated texts, e.g., ironical or highly metaphorical writing.

There is also another thread of thought about topics starting from the formula *about X*. As I mentioned before, the kind of expression filling the variable in the *about*-formula is mostly an NP, like in *about the coming elections*, and expressions of the category NP are standardly used to refer to entities of vastly different kinds, from concrete objects like (individual) trees, persons and events to abstract objects like ideas and propositions.[4] All these things can be talked about. Thus, seen from this point of view, a topic is an entity to be talked about or "elaborated on". From this assumption we can move on to an interesting aspect of our knowledge: Our knowledge is, at least partly, organised by objects. We have – various amounts of – knowledge about trees, certain persons and events, and even knowledge about some ideas and propositions. This kind of knowledge, which sometimes is and sometimes may become common knowledge, is an important resource for our talking and writing about given topics.

Another hopeful approach to understanding the status of topics could consist in claiming that a topic provides a criterion of *relevance* for the contributions to the respective discourse.[5] Therefore, if a speaker, for instance, professes to talk about X they should be prepared to justify individual statements

---

[3] A brief critical overview of methods of topic detection and topic-based text segmentation is given in Stede (2012, ch. 2). See also recent work on topic modeling for short texts, e.g. Twitter tweets, (see Qiang et al. 2022) and the clustering of collections of such texts according to topic (see Bilal et al. 2021).

[4] A classic source inspiring reflection about what kinds of entities we think and talk about is Strawson's "Individuals" (Strawson 1959).

[5] Cf. the "Principle of Relevance" Strawson introduced in his reflections on aboutness (Strawson 1964/1971, 92f.).

made in the course of their talk by showing their relevance to the topic, maybe not for all statements, but at least for a majority of statements or such statements as they particularly emphasized, or maybe individual paragraphs. This might still be considered somewhat vague, but it has the advantage of at least hinting at a procedure for deciding the question if the talk was "on topic" or not.

It also turns our attention to a point mentioned before: Topics are not present in the surface of a text, i.e., the syntax and the lexical elements, although they may be *indicated* by "topic expressions". They are an aspect of the *meaning*, in the sense of *speaker meaning*, or the *understanding* of a text. Thus, talking *about* – and therefore the concept of topic – is a *pragmatic* concept, like referring, and not a semantic concept. This parallel to the practice of referring will hopefully become more clear in the following sections of this chapter.

A related approach to further clarification could consist in reflecting on the function a topic has for a speaker or writer. Generally speaking, we could say that topics provide a framework for what can or should be said in a discourse. This is, of course, related to the idea of a criterion of relevance. Thus, topics seem to have a *control function* for the production of contributions to discourses, in other words they are an aspect of a communication principle. In applying this kind of principle, speakers and writers – and analogously hearers and readers – use various resources, e.g., their knowledge of topic structures, their knowledge of types of topic development, their knowledge of the relevant text type, their knowledge of the common ground reached at a given point in the discourse etc. So, in a way, a topic with its related knowledge should in itself be considered a communicative resource. I shall deal with the question of how participants of a discourse draw on this resource in the course of this chapter.

The control function I mentioned just now has a further aspect worth mentioning. Topics not only determine – to a certain degree – what can or should be said, but they also give orientation as to what can be said in which *sequential order*. In other words, a topic is also "a sequential structuring mechanism" (Goutsos 1997, 29). This is partly due to the fact that part of the topic knowledge mentioned above is knowledge about typical sequences of topics and subtopics that we acquire as readers and as participants in conversations. Parents of small children, for instance, know that in talking about children's illnesses topics like the following are standard and also occur in this and similar sequences: the current illness(es), symptoms, fever, possible therapies, good (and bad) doctors, parents' stress caused by the illness and the fact that children cannot go to daycare etc. One should, however, not consider this "mechanism" to be too restrictive, as it is always possible to improvise on a topic.

Both the aspects mentioned before, what can be said (or written) and how one subtopic may follow another, can be considered aspects of *topic management*. Topic management includes activities like introducing a topic, announcing a topic, organising topic continuity, changing the topic, making explicit the topic of a certain passage of discourse etc. There is much to be said for approaching the role of topics by analysing methods of topic management. In the first place, discussing the *practice* of topic management helps to avoid some of the pitfalls involved in asking what topics (really) are.[6] Secondly, this practice provides a vivid illustration of what the point of having "topics" at all consists in. And thirdly, this view also helps to avoid a too static view of topics. In many cases the topic of a text or dialogue is a dynamic achievement that emerges in the flow of utterances by means of small steps, some of which are not foreseeable at all, so that it is only at the end of the discourse that one really knows what the topic is – if at all.[7]

Finally, an important theoretical observation should be mentioned at this point. Topics as factors of the coherence of texts/conversations interact with other sequencing mechanisms, e.g., sequencing patterns, coreference, and common ground management. I shall confine myself here to giving a few examples of the interaction of topics and sequencing patterns. For instance, a series of *statements* which otherwise do not show a strong functional utterance-to-utterance connection can be made coherent by the fact that the whole ensemble is meant and understood as a *description* and the fact that all these statements contribute to the same topic. Thus, the sequencing type "description" and the topic interact. Similarly, sequences of statements about a past event can be functionally connected by being part of a *narrative* and thematically by being *about* the same event. This pattern is quite frequent with news reports, as Schröder (2003, 219) noted in his analysis of text types in newspapers. A final example, which I shall go into in more detail in section 14.3, concerns sequencing patterns which are sometimes used to manage topic shifts, e.g., explanations leading from the thing to be explained to explanatory factors that are topically (more or less) distant from the starting point of the explanation. Thus, the topic shift interacts with the sequencing pattern of "explanation".

Summing up this section one could say: (i) A topic is an object about which one can talk or write (or think). (ii) Topics are not expressions and are not "in" the text. (iii) Topics provide a criterion of relevance for contributions to a

---

[6] I discussed some of these pitfalls in Fritz (2017, ch. 4), but I shall not go into the history of this problem here.

[7] It is worth remembering in this context that, as Sacks put it, in good conversation "new topics are never 'introduced' they just happen along" (Sacks 1995, vol. II, 352). For the stepwise movement of topics in dialogue, see also chapter 14.

discourse. Thereby they function as a kind of control mechanism for what is to be said or written. This "mechanism" is, however, not always very restrictive. (iv) Topic-related knowledge provides resources for "topic talk" and the sequencing of topics in discourse. (v) The complex practice of dealing with topics in discourse can be called "topic management". Topics frequently only emerge in the course of this practice. (vi) Topics interact with other sequencing mechanisms (sequencing patterns, coreference, common ground management) in creating the coherence of a series of utterances.

So far, there is still one important feature of topics missing, the fact that they are structured in different ways. I shall therefore, in the following sections, start with some observations on the structures of topics and from there go on to present and discuss forms and strategies of topic management.

5.2 Topic structures

Topics can be analysed in terms of (at least) four types of structures, which play a role for various aspects of topic management and the topical coherence of discourse. The four types of structures I shall discuss are the following:

(i) network structures,
(ii) hierarchical structures,
(iii) topic structures with *by*-relations,
(iv) linear (or sequential) structures.

In using this approach, we should, however, bear in mind that topics tend to be "open textured" and in a permanent state of flux. In other words, the primary empirical fact is the flow of "topic talk" in various groups, contexts and media of a society, from which these forms of structural organisation emerge. It is therefore necessary to differentiate between pre-existing "public" topics, which can be used as resources in topic talk, and individual topics achieved in the course of an individual discourse. The following sections will start by dealing with public topics and their associated topic knowledge.

5.2.1 Network structures

When looking more closely at topics with a minimal degree of complexity, we soon find internal structures that can be represented in form of network graphs. I shall start by having a look at a fairly recent topic, namely Covid, which only emerged in 2020 and which could be considered common knowledge in 2020/2022.

What this example shows, as an extra insight, is that topics and topic networks are *dynamic*. They emerge, sometimes being explicitly introduced or supported by certain agents, they grow in complexity, they change and sometimes they fade away, and they are open in many directions.[8] Sometimes certain subtopics become more prominent at a certain time, e.g., the role of different vaccines, or new subtopics arise, like the topic of different "strains" of the virus and the danger they may cause. And, of course, each individual and each community (e.g., each family or "filter bubble") possesses an individual network of topics, which is built up in the course of their particular communication history and which relies in varying degrees on general knowledge. Frequently public topics arise in the course of controversies conducted in public media, and some topics become characteristic of discourses in certain groups.

Another fundamental point is that for each element in such a network the speakers/writers and their audiences have at their command a certain, if varying, amount of knowledge that can be used in communication and that contributes to the creation and perception of coherence in discourse.[9]

To give a first idea of what is involved in such a topic, I shall now give a selected list of (sub)topics that somehow "belong" to the topic of Covid:

(1) the Covid pandemic
(2) the coronavirus (Sars-CoV-19)
(3) the disease (COVID-19)
(4) symptoms of the disease
(5) mortality and long-time effects of the disease
(6) therapies
(7) testing
(8) vaccines
(9) hospitals
(10) intensive care wards
(11) ventilators
(12) social distancing
(13) face masks
(14) lockdown
(15) closing schools
(16) consequences of lockdown for the economy
(17) consequences of lockdown for families
(18) protest against preventive measures

Looking at these topics, which everyone will be familiar with who read newspapers or social media or talked to a number of people in the course of

[8] It is not surprising that the Covid topic was (partially) superseded in public media by the topic of the Russian invasion of Ukraine after February 24th, 2022.

[9] In Cognitive Linguistics knowledge structures of this kind are discussed in terms of "frames" (see, for instance, Ziem 2014).

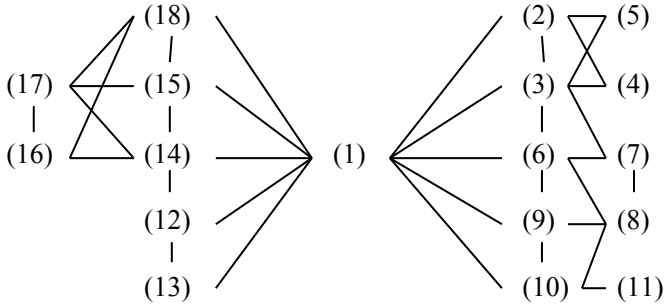
2020/22, we find that some seem to be more closely related to others and some seem to form clusters. For instance, the topic of lockdown seems to be closer to the topic of the consequences of lockdown than to the topic of symptoms of the disease. This means, in practice, that it is easier to move directly from one (sub-)topic to the other in one case than in the other.[10] In the graphic presentation below this would be represented as a shorter "path" between the vertices (14) and (16) than between (14) an (4). And, as for clusters, one finds a cluster of topics surrounding the disease as such, e.g., the virus that causes the disease, its symptoms, mortality and (long-term) effects. Further clusters are medical activities, including testing and vaccination, and government-imposed prevention measures (social distancing, lockdown, obligatory face masks). It is quite normal that individual topics can be assigned to different clusters, e.g., testing belongs both to medical activities and to government-imposed prevention measures.

This kind of cluster-formation is in part a result of how these (sub-)topics are discussed in the media and in personal communication. Talking about prevention measures one can easily move from one to the other, e.g., in the sequence "social distancing" – "face masks" – "lockdown". From this cluster there is again an easy "bridge" to topics like the consequences of prevention methods, and from there again to protests about such measures, and so on. On the other hand, the practice of talking about these topics by moving from one member of the cluster to the other stabilises the cluster. Generally speaking, we can say that topic networks and the associated topic knowledge facilitate "topic talk", including the coherent sequencing of topics, and, conversely, that the practice of topic talk causes the emergence of topic networks.

Of course, not all topics raised in communication are already members of established networks. New topics are continuously being introduced, in everyday conversations, in the sciences and in the arts.

In addition to the observation of clusters we can also see connections between the clusters, e.g., a close connection between the disease cluster, its symptoms and effects and the "hospital" cluster. Or the "social prevention" cluster and the "medical prevention" cluster, including testing and vaccines. (This connection is not shown in the illustration below.) Trying to illustrate this structure by a (very much) simplified network graph – a so-called undirected graph – we would get something like the following picture:

[10] Of course, there are "jokers" like *that reminds me of*, which facilitate long-distance topic jumps.



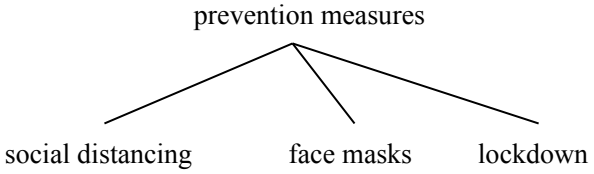
This kind of structure could, in principle, be shown for any topic that already had a certain life span. It is worth noting that this network representation is intended as an "object of comparison" which is meant to show similarities and dissimilarities to our actual practice of talking about these topics (see Wittgenstein 1958, §§ 81 and 130). Maybe a multidimensional mesh would make a more realistic object of comparison. But such a multidimensional rhizome-like structure is difficult to represent graphically (for the concept of rhizome, see Deleuze/Guattari 1977).

Topic networks constructed by an algorithmic analysis of a given text corpus have a different empirical status: (i) They are mostly based on the assumption that topics are represented by surface topic expressions – which is not trivial, as I showed –, and (ii) they represent collocation frequencies of relevant expressions. Such analyses provide interesting, albeit indirect data for the interpretation of network structures and their dynamic (see, for instance, Bubenhofer et al. 2015).

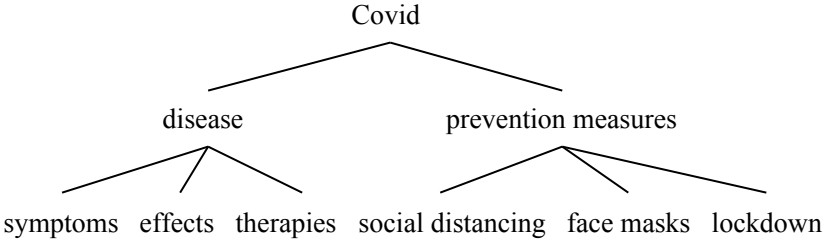
The point to take away from this illustration is that speakers/writers and their audiences have available knowledge structures of this kind as a resource for participating in discourses. Such networks of topic knowledge can be quite different with different people, but, fundamentally, every speaker/writer presupposes their audience to participate in some shared knowledge of this kind for them to rely on in performing their topic management.

5.2.2 Hierarchical structures

Over and above network structures, topics can be seen to form hierarchical structures. In the example looked at just now we find that a general topic like prevention measures can be discussed by discussing specific subtopics, e.g., social distancing, face masks, and lockdown. This fact could be presented in form of a tree representing a hierarchical topic structure:



Adding another level to this kind of hierarchy, we can produce the following tree:



Structures of this type are frequently used for the global organisation of topics. The top level, for instance, could be the topic of the whole text, whereas the next level represents chapters and the lowest level represents paragraphs within the chapters. In scholarly writing, this kind of topic structure is sometimes represented in form of a decimal system like the following:[11]

Covid

1. Disease
1.1 Symptoms
1.2 Effects
1.3 Therapies

2. Prevention measures
2.1 Social distancing
2.2 Face masks
2.3 Lockdown

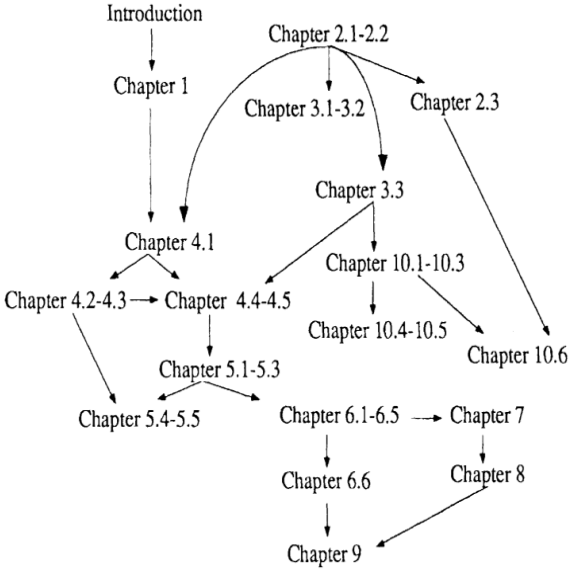
In the tree representation the clusters of the network reappear as the members of a branch of the tree, however with an additional hierarchical structure. In the practice of scientific writing, the step of organising one's topics in such a fashion that a hierarchical structure becomes apparent is frequently an

[11] An early example of this form of representation of the topic structure is to be found in Wittgenstein's *Tractatus logico-philosophicus* (1921).

important step in ordering one's text. This is a step that is often difficult for beginners, e.g., students starting to study a new and unfamiliar subject.

A well-known type of structure that contributes to hierarchical topic structure is represented by so-called mereological structures, i.e., structures that – in various guises – consist of the relation of part to whole. As examples of this (fairly heterogeneous) type of structure one could name everyday relations like twigs being parts of trees or bungalows being houses, and houses again being buildings. The latter type of structure is frequently used in scientific classifications (taxonomies), e.g., in ornithology, where chaffinches, bullfinches and goldfinches are classed together as finches, and finches, together with warblers, thrushes and – strangely enough – ravens, are classed as songbirds. From the point of view of topics this means, that mereological structures can be used as hierarchical topic structures, for instance, if we write an essay about finches, dealing with chaffinches, crossbills and goldfinches. Or, to change the perspective, mereological structures are really a kind of topic structure.

A sophisticated example of an author's reflection of the topic structure of his book is provided by Asher's introduction to *Reference to abstract objects in discourse* (Asher 1993, 12f.). After explaining the book's topic organisation, he introduces the dependency diagram shown below, which represents the connection between the organisation in chapters (topics in decimal notation) and the dependency structure of these chapters, which also indicates a (partial) sequential relation between chapters and sub-chapters.



Without presupposing any knowledge of the topics referred to, I shall give a brief example explaining the structure of this representation, starting from Chapter 4.1 on the left side of the diagram ("The semantic representation for sentential nominals"). By the arrows leading to this chapter, we are shown that the chapter presupposes knowledge of Chapter 1 ("Abstract entities and the nominals that denote them") and Chapters 2.1 and 2.2 ("A crash course in DRT").[12] Therefore the reader could either follow the numerical sequence of the chapters 1 – 2.1.-2.2. – 4.1, or, if they know the author's position on DRT, they might – at their own risk – omit 2.1.-2.2 and continue directly with Chapter 4.1. But, of course, the reader can also strictly follow the numerical sequence from Chapter 1 to Chapter 10, a reading strategy which the author might prefer. Later on, these dependencies become more complicated, but, anyway, the reader gets a vivid idea of connections over and above the hierarchical structure indicated by the decimal notation.

Another interesting example of an overview of a book in form of a network of topics is the network of "concepts and lines of enquiry" and "links emphasized by different authors in this book" presented in the introduction of the collected volume of papers on "Roots of human sociality" (Enfield/Levinson 2006, 10).

5.2.3 Topic structures with *by*-relations

In the chapter on basic elements of texts I showed that speech acts and activity types in general can be analysed to have an internal structure, of which the *by*-relation is an aspect. One can, for instance, perform an accusation *by* making a statement (with additional commitments). A similar kind of structure can also be found in relation to topics. One can talk or write about topic_1 *by* talking or writing about topic_2. Of course, not any pair of topics will do for the production of such a structure. In the following I shall mention two related types of *by*-relation topic structures.

The first relevant type of structure has already been implicitly mentioned in the previous section on hierarchical structures. A writer could, for instance, write an essay on songbirds *by* writing about finches, thrushes, and nightingales, even if he never explicitly uses the word *songbird*. A reader asked about the topic of the essay could give an appropriate description of the essay by saying it was *about songbirds*. Thus, a selection of elements from the lower levels of a mereological structure can be used as a *by*-relation topic for the higher category. This topic strategy however presupposes some extra

[12] DRT is short for Discourse Representation Theory.

knowledge for writers and readers: They have to know that these birds are songbirds, which is a fairly simple kind of knowledge. For readers who do not know that ravens are songbirds, it might however seem doubtful if an essay on blackbirds and ravens is *about songbirds* – they could think, for instance, that it was *about black birds*.

A second type can be considered a special case of the former. This type could be called a *pars-pro-toto* topic.[13] In this case one can talk/write about topic$_1$ *by* talking/writing about topic$_2$, if the latter is a particularly *salient* or representative subtopic of the first topic. A case in point is a lecturer who gives a lecture on sequencing patterns for texts. A student asked about this lecture might correctly answer: She talked *about coherence of texts*, because, according to some authors, sequencing patterns form one of the basic factors of text coherence.

5.2.4 Linear (or sequential) topic structures

As I already noted in connection with the example of the complex topic structure of a scholarly book mentioned above, texts – as well as conversations – also have a linear or sequential topic structure, which means that the topics (and subtopics) of a coherent discourse do not follow each other in any old way, but on the basis of certain sequencing principles. Arguably, the linear structures are the fundamental structures of topics, because it is in their sequence in discourse that we encounter them in the first place, and dealing with topic sequencing is one of the fundamental tasks of topic management. Some of the principles for topic sequencing are very general, others depend on the respective text type. In the following passages I shall give examples of the first type. Text-type specific sequencing principles will be discussed in the chapter on topic management (5.5.3). Generally speaking, there are no mechanical rules for sequential topic management that apply to all kinds of discourse. But for discourses where comprehensibility and usability are at a premium, a number of rough principles can be named.

One of the basic sequencing principles I mentioned earlier on in connection with knowledge management is the principle of systematic knowledge building. This principle says one should only present a piece of knowledge K_2 if the knowledge K_1 necessary for comprehending the presentation of K_2 has either been presented earlier on in the discourse or can otherwise be considered common knowledge. Therefore, a coherent sequence should be $K_1 > K_2$ – and not $K_2 > K_1$. This principle has a direct application to the sequencing of topics. If

[13] In classical rhetoric pars pro toto ('a part for the whole thing') is a rhetorical figure. An example: *The sails are coming in* meaning *The boats are coming in*.

introducing topic T_2 in a certain portion of a text presupposes knowledge connected with T_1, one should either first deal with T_1 or make sure that the knowledge relevant for the comprehension of T_2 is already common knowledge for other reasons, e.g., in case of general background knowledge. This simple principle is, for instance, fundamental for teaching discourses, it is, however, frequently violated, which tends to lead to grave comprehension problems.

More detailed principles, which partly rely on this principle, can be given as follows:

(i) Closely related topics (subtopics) should be placed in close sequential neighbourship.
(ii) In case of a discontinuous placement of closely related topics cross-references should be given.
(iii) Topics belonging to the same hierarchy level should be dealt with in close sequential neighbourship.
(iv) Changes from one hierarchy level to the other should be explicitly indicated.
(v) When linearising topics one should follow well-established topic connections – a variant of (i).
(vi) If you introduce topics in a certain order, explore them in the same order.

As is usually the case with communication principles, it frequently happens that they come in conflict with each other. It sometimes occurs, for instance, that an author has to deal with two closely related topics T_1 and T_2, which should therefore be placed in immediate sequential neighbourship. However, introducing T_2 presupposes knowledge that has so far not been provided in the text and can also not be considered common knowledge. In this case, following the principle of systematic knowledge building, the author will have to interrupt the topic sequence $T_1 > T_2$ by introducing a text passage giving the necessary background knowledge for T_2. To mitigate the emerging problem of coherence, the author could bridge the gap by explicitly indicating his digression from the desirable sequence of topics. An alternative solution to this conflict of principles could consist in providing the information necessary for the comprehension of T_2 at some other, earlier point in the text. This example shows that sometimes "local" sequencing problems require "global" solutions, a situation that experienced writers are usually well aware of. It also shows why problems of topic management can sometimes be quite difficult.

5.3 The vagueness of topics

Before leaving the topic of the structure of topics I want to correct a misconception I might have fostered in the course of my analysis so far. Many topics are not as clear-cut as it might seem and as the nodes in my network graph might suggest. Many topics obviously have fuzzy edges and in many cases subtopics "belong" to different topics. Therefore, transition from one topic to the next may proceed by a technique of topic management that has been called "topic shading", i.e., by subtly and imperceptibly moving from one topic to the next. Talking, for instance, about sadness, melancholy, and depression, we might be in doubt how these topics are related to each other and where one ends and the other begins, as happened to Robert Burton, who, in his book *The anatomy of melancholy* (1621), complained about the "perplexities" involved in analysing the different species of melancholy (Burton 1964/1621, vol. I, 177).

And where concepts have a family-resemblance structure like in the case of Wittgenstein's games example (board games, card games, ball games etc.), the corresponding topics share this structure.[14] The same is true of categories with a prototype structure, e.g., the category of birds. If somebody reports having listened to a conversation about birds (feathered), we would probably be surprised to hear that the persons involved had mainly talked about different kinds of penguins, a rather marginal kind of bird, and not about robins and finches.

Examples of (sub)topics that belong to two or more different topics are also not difficult to find. If, for instance, we talk about cherry trees in general and about the trees in our garden, these seem to be distinct, if somehow related, topics. We could, however, in some contexts, move quite inconspicuously from one topic to the other by talking about the cherry tree in our garden. Thus, the structures of topics provide many opportunities to walk the edges of topics and to move imperceptibly from one to the other.

5.4 Doubtful topics

Another point worth remembering concerns the situation where it is doubtful what the topic really is. For most of the examples discussed in the sections of this chapter it can be assumed that it is fairly clear what the topic of a given discourse is. But this is not generally the case. Sometimes, topics are not only vague, so that it is not clear which of a number of related topics is actually being talked or written about, but in some cases the topic is hidden on purpose

[14] See Wittgenstein (1958, § 66).

or potential recipients are deliberately mislead as to what the topic really is. Such strategies can also cause problems concerning the perception of the coherence of (parts of) a discourse.

One type of doubtful topics are taboo topics, i.e., topics that are not accepted and therefore to be avoided in certain contexts. Depending on the respective culture and context, certain topics like (aspects of) sex, aspects of religion or social relations can be taboo in this sense. In cases where writers or speakers feel they have to deal with such topics despite the taboo they sometimes use *euphemisms*, i.e., expressions that help to avoid the typical topic expressions while allowing the speaker/writers to refer to the problematic entity. In medieval religious writings, for instance, writers sometimes tried to avoid the name of Satan and used circumlocutions instead, e.g., *antiquissimus seductor* 'the ancient seductor', *pietatis inimicus* 'the foe of piety', or *ueritatis hostis* 'the enemy of truth'. Euphemisms are also frequently used in business relations. Receiving a letter from our insurance company with the headline *Adjustment of premiums* we will not be fooled: We know the letter is about raising the insurance premiums once more. A related phenomenon is the practice of parents to hide certain topics from their children ("not in front of the children") by, for instance, talking of uncooperative people in general and meaning the lack of cooperation of their neighbours or colleagues.

Another situation where the topic of a text may be doubtful is the use of irony. Giora focused this problem as follows: "Indeed, irony may be lost on some audiences, who would miss the theme of the piece if their search for an appropriate interpretation stops at the literal level. However, if read as an ironic piece, the theme of the novel makes a lot more sense, and its social implications more relevant" (Giora 2002, 291).

To conclude this small selection of doubtful topics I give a literary example. A literary technique of (partially) hiding a topic is the *roman à clef*, i.e., a novel in which real people are depicted under fictitious names. In one passage of his book *The man in the red cloak* (2019) Julian Barnes writes about the novel *A Rebours* by the French author Joris-Karl Huysmans (published in 1884), which by some people was considered to be a *roman à clef* about the aristocrat Robert de Montesquiou (1855-1921). On this (mis)perception, Barnes writes:

> *A Rebours* isn't 'really about' Robert de Montesquiou. Huysmans spins off from a few pillaged details about the Count's ménage, but then focuses on his own thematic obsessions (Barnes 2019, 188).

Thus, we see that in some cases it is not clear what the topic of a discourse really is, which might lead to debates about the "real topic".

5.5 Topic management

5.5.1 Some general remarks on topic management

Topic management as it is discussed in this chapter comprises quite an array of tasks, which I shall briefly outline in this section. The two main tasks consist in *organising* the topics of a text and *indicating* the organisation. Both these tasks are closely related to the task of creating and preserving coherence.

Organising topics includes creating and structuring thematic connections on the one hand and bringing them into linear order on the other. In doing this, the speaker/writer also has to embed the topics and their connections in the functional structure of the text, e.g., a description, an argumentation, or a narrative. At the same time, the speaker/writer has to monitor the process of common ground building. As I showed in the preceding sections, topics and their relations have various structural properties, which have to be taken into account when organising the topics of one's text. What is usually the most difficult part of this task is the linearisation of the topics and subtopics in the flow of the text.

Basically, there are two types of topic development, namely topic continuity and topic change. Whereas topic continuity is frequently not marked explicitly, as it can be considered the default mode, topic change is often indicated by various means, of which I shall mention a few presently. But topic change can also be performed without explicit indication, for instance in cases where common knowledge of the structure of the respective topics and subtopics can be assumed.

Generally speaking, indicating the organisation of topics explicitly is a task which is important for the speaker/writer's own orientation in their text and which, on the other hand, is a fundamental aspect of "recipient design". For the purpose of indicating topics and topic structures in writing, a wide array of means has been developed, of which some already have been mentioned and used in this chapter, e.g., tables of contents, introductory passages, headlines, marginalia, hashtags, bullet points, decimal notation, indentation, emphasising key words, placing topic words at the beginning of a paragraph, cross-references etc.

5.5.2 Local and global topic management

An important unit for topic continuity and change is the paragraph, as I mentioned earlier on in my introduction of the basic text units. Within a paragraph

a topic or subtopic can be established and elaborated, whereas moving from one subtopic to the next is performed by creating new paragraphs. Seen from the point of view of the recipient, indicating a new paragraph helps the reader to understand the topic movement from one subtopic to the next.

To give an example: In an article explaining some of the basic ideas of genetics, Richard Dawkins first introduces the Human Genome Project (HGP) and then goes on to discuss some aspects of the Human Genome Diversity Project (HGDP). The following paragraphs show some of the "work" performed in introducing and developing these topics:

> By contrast the Human Genome Diversity Project (HGDP), now under way, builds on the foundation of the HGP but focuses on those relatively few nucleotide sites that vary from person to person, and from group to group. […][15]
>
> The medical benefits of studying human variation could be immense. Hitherto, almost all medical prescribing has assumed that patients are very much the same, and that every disease has an optimal recommended cure. Doctors of tomorrow will be more like vets in this respect. Doctors have only one species of patient, but in future they will subdivide that species by genotype, as a vet subdivides his patients by species. […] The point is that for some diseases there may be as many different optimal treatments as there are different genotypes at a locus – more even, because generic loci may interact to affect susceptibility to disease.
>
> Another important use of the genetics of human diversity is forensic. […] (Dawkins 2003, 31)

In the first paragraph Dawkins introduces the general topic of research on human genome diversity by pointing out the HGDP, contrasting it with the HGP described before, and briefly explaining the particular focus of this project. He then goes on to describe the potential practical value of this kind of research – which forms a subtopic of the general topic –, introducing as a first subtopic of this topic its medical benefits. In order to indicate this subtopic, he uses an indent and opens up the first sentence of the new paragraph with the topic-indicating expression *the medical benefits of studying human variation*. This topic is then locally developed by showing that for different genotypes there could be different specific medical treatments. The following paragraph deals with a further subtopic ("another important use of the genetics of human diversity"), namely forensic use, which is explicitly introduced by the corresponding thematic expression *Another important use.*

So what we see here is the combination of local topic management *within* paragraphs and global topic management in a series of paragraphs developing subtopics of the general topic.

---

[15] A nucleotide is a constituent of DNA.

5.5.3 Topic organisation in different text types

In the course of the history of discourses various text types have evolved which embody the experience of different communities gained in performing various communicative tasks. For many text types the rules or principles of topic organisation to be followed by writers are fairly vague and flexible, like in the case of private letters, diaries, feature articles, sermons, lectures, and novels. In other cases, clear models and rules have evolved, based on legal or usability requirements and other functional constraints. Examples of such text types are examination regulations, medical leaflets for patient information, scholarly research articles, and news articles. For these text types, *criteria of global coherence* are given by these models and rules, which concern both the functional elements and the topics, including their linear organisation.

In the present section I shall give two examples of such models and rules for topic management, namely *medical leaflets* and *scholarly research articles*. I shall try to show for these text types that topic structure and topic management are fundamental factors of discourse coherence and that, therefore, knowledge of relevant topic structures is a basic resource for producing and understanding coherent texts of these types. At the same time, I shall indicate how the standardisation of such text types develops historically. As I shall go into some detail in doing this, readers may need a bit of patience.

Starting with medical package leaflets, I first identify a basic communicative task to be solved by such leaflets, which has quite a long history: Persons using medicinal concoctions should know what a bottle of medicine, for instance, contains, what it should be taken against, and how much should be taken. The practice of using slips of paper fixed to medicine bottles for this information goes back at least to the time around 1800, probably longer. In the course of the 20th century, two factors seem to have been particularly influential in promoting the development of regulated forms of package leaflets, namely liability issues and the information needs and safety of patients. As this development was mainly a matter of national policy, the form and speed of development was different in different countries. In the following I shall start with a remark on the history of package leaflets in Germany and then go on to describe the situation in the EU. The German Medicine Act ("Arzneimittelgesetz") of 1961 prescribed some general information to be placed on the outside of the container of the medicine, including the name of the firm, the medical composition, the form of application, and the expiry date. This information was mainly meant for doctors and pharmacists and it was not placed inside the respective package. Partly due to liability issues raised in legal processes during the 1960s and 1970s, various initiatives from doctors and the pharmaceutical industry suggested guidelines for the introduction and contents

of package leaflets, which were incorporated in a paragraph of the new German Medicine Act of 1976 (put into effect in 1978). The package leaflet was required to contain information relevant for the user, e.g., therapeutic indications, contraindications, undesirable side effects, and interactions with other medicines. The exact sequence of these elements was not prescribed at the time and was only made obligatory in 1994.

Mainly for legal reasons, the form of "medicinal product information" was later also strictly regulated in the EU (1992). This regulation was implemented by requiring a "Summary of Product Characteristics (SmPC)" to be included as part of the application for the market authorisation of a medicinal product.[16] The topics to be covered in this summary were and still are strictly prescribed. As special "Information for the patient" a 'package leaflet' is to be included, for which a template is given. The "headings and standard statements given in the template must be used whenever they are applicable". As one of the elements of the leaflet a "contents listing" is to be provided. This list, formulated in language that the users are expected to understand, follows the necessary topics and their prescribed sequence:

1. What X is and what it is used for
2. What you need to know before you <take> <use> X
3. How to <take> <use> X
4. Possible side effects
5. How to store X
6. Contents of the pack and other information

In the template a second level in the hierarchy of topics is required. With this second level, which is noted in square brackets and which is not in "user language", the topic structure is as follows:

1. What X is and what it is used for
1.1.[(Invented) name, active substance(s) and pharmacotherapeutic group]
1.2.[Therapeutic indications]
1.3 [Information on the benefits of using this medicine]

2. What you need to know before you <take> <use> X
2.1 [Contraindications]
2.2 [Appropriate precautions for use; special warnings]
2.3 [Interactions with other medicines]
2.4.[Interactions with food and drink]
2.5 [Use by pregnant or breast-feeding women, information on fertility]
2.6 [Effects on the ability to drive or to use machines]
2.7 [Excipients warnings]

[16] See https://www.ema.europa.eu/en/documents/template-form/qrd-product-information-annotated-template-english-version-102-rev1-highlighted_en.pdf.

3. How to <take> <use> X
3.1 [Dose]
3.2 [Route(s) and/or method of administration]
3.3 [Duration of treatment]

4. Possible side effects

5. How to store X
5.1 [Expiry date]
5.2 [Storage conditions]
5.3 [Where applicable, shelf life after reconstitution, dilution or after first opening the container]

6. Contents of the pack and other information
6.1 [Full statement of the active substance(s) and excipient(s)]
[…]

From the point of view of the coherence of package leaflets, the many specifications in section 2 ("What you need to know before you take X") are particularly interesting. These topics concern possible risks for the user and possible legal problems for the producer, so they form a cluster of related topics. Functionally these subdivisions are warnings and are therefore placed early on in the text.

In early package leaflets the "Full statement of the active substances and excipients" (6.1) was frequently placed at the very beginning of the text, after the name of the product. As users mostly did not understand this chemical information anyway and (therefore) were not really interested in it, they tended to give up reading the leaflet at this early point. Usability studies discovered this problem and suggested putting this information at the very end of the text. So in this case topic sequencing was dictated by usability. There were also other usability problems concerning the use of *topic-indicating* expressions like *contraindications* and *interactions*, the exact meaning of which non-specialists frequently did not know. These comprehension problems were also discussed in research on usability of medical information and were later taken into account in the wording of the individual sections for non-specialist users.

The advantages of such a standardisation of the text type for the user mainly consist in the predictability of the individual elements and the habitualisation of users to these elements and their sequence. Thus, the standardised template provides a criterion of coherence both for the producer and the user.

As mentioned at the beginning of this section, the second text type I shall discuss here is the *scientific research article* or "scientific paper", as it is normally called today.[17] The history of the "modern" research article goes back

[17] For the history of the research article, see Bazerman (1988), Atkinson (1999) and Gross/Harmon/Reidy (2002).

to the late 17th century, when the first scientific journals were established, the "Journal des Sçavans" (1665), the "Philosophical Transactions of the Royal Society of London" (1665), and the "Acta Eruditorum" (1682).

In these journals there was much more variety of text types than what we are used to from most scientific journals today. There were short reports of extraordinary occurrences like the appearance of a comet or the birth of a deformed calf, there were extracts of letters to the editor, reporting scientific experiments, descriptions of scientific instruments (e.g. microscopes and telescopes), short theoretical treatises (e.g. on the topic of calculus), contributions to controversies, reviews of books etc.[18] So the rules for the kind of text to be included in the journal were quite flexible, as long as the texts could be said to contribute to the "spreading of useful knowledge", "the further discoveries of nature", and "philosophical observations" in general. Even a hundred years later, when the professionalisation of sciences like chemistry had so far progressed that specialist journals for different disciplines began to appear, the kinds of texts to be found there were quite varied, including letters to the editor, narratives of experiments, descriptions of instruments, theoretical disquisitions, and contributions to controversies.[19] This process of professionalisation and specialisation has continued up to the present day and has spawned various text types or genres of scientific publication, up to the latest phase in this evolution, the digitalisation of scientific publishing, which has produced new text types using new forms of multimodal representation (e.g. new types of graphs), links to data bases and related publications etc.

The most well-known research article format in the experimental sciences, the so-called IMRAD format, evolved in the course of the twentieth century and became a standard format in the 1970s.[20] *IMRAD* is an acronym for I*ntroduction*, M*ethods*, R*esults*, *a*nd D*iscussion*, the headings of the main sections of an original research article of this format. One aspect of the coherence of texts of this format consists in the fact that these headings reflect aspects of the standard practice and the "rhetoric" of (certain) experimental sciences, so that readers and writers familiar with this practice could use their familiarity

---

[18] An interesting summary of the kinds of topics dealt with in the first four years of the "Philosophical Transactions" was given in a preface at the beginning of the fifth year (No. 45, 1669), from which I have taken the phrases quoted in the following sentence (https://royalsocietypublishing.org/doi/pdf/10.1098/rstl.1669.0002).

[19] On the use of journal articles in a 1790s chemical controversy, see Fritz (2018b).

[20] For the evolution of the experimental report in physics (1893-1980), see Bazerman (1984); for the history of the use of IMRAD in medical journals from 1935 to 1980, see Sollaci/Perreira (2004).

as a resource for understanding and producing papers of this type.[21]

One of the advantages of this format, according to some authors, is that the reader can easily navigate the text and find relevant information:

> The IMRAD structure facilitates modular reading, because readers usually do not read in a linear way but browse in each section of the article, looking for specific information, which is normally found in preestablished areas of the paper (Sollaci/Perreira 2004, 366).

Using this format is nowadays not strictly required for the publication of a paper, but journals indicate in their guidelines for authors that it is "usually" used or "useful", which certainly has a normative effect, even if the wording is not explicitly prescriptive:

> The text of articles reporting original research is usually divided into Introduction, Methods, Results, and Discussion sections. (ICMJE Recommendations 2019, 14)

> Many authors find it useful to organize their manuscripts with the following order of sections; title, author line and affiliations, keywords, abstract, significance statement, introduction, results, discussion, materials and methods, acknowledgments, and references. Other orders and headings are permitted. (PNAS template)

As can be seen in the template for the Proceedings of the National Academy of Sciences (PNAS) and in many journals, a (structured) abstract has become an additional obligatory feature of the article structure. As the abstract is frequently a condensed version of the IMRAD structure, it plays an important role in the general topic management of research papers, both for human readers and digital search engines.

From the point of view of topic management, there is frequently a particularly close topical connection between the title of the paper, where the main topic is first presented, the abstract, the introduction, where the topic is introduced in more detail and (often) related to other, higher order topics, and the central "results" section, where the topic is elaborated in detail. This close topical connection is an important factor of the coherence of a paper written in this format.[22]

As in other text types or genres, the global coherence of the text relies on both topical and functional (action type) properties. I shall therefore add a few remarks on the functional (or act) structure of the IMRAD format.

---

[21] Bazerman (1987) showed that the publication manual of the American Psychological Association reflected a behaviourist view of psychological research.

[22] It is worth noting that present-day guidelines frequently place the "methods" section last in the sequence of IMRAD elements. In some journals it is presented in small print, which also indicates reduced importance.

Interestingly, the four headings of the IMRAD schema fall into two categories, namely expressions referring to action types and expressions referring to topics. Giving an introduction and leading a discussion are action types, whereas methods and results are abstract objects which can be referred to as topics and which, from a functional point of view, are mainly dealt with by giving descriptions. Whereas the function of introductions consists mainly in providing background knowledge for placing the study and its topic in its research context, discussions can, among other things, be used to "emphasize the new and important aspects of the study and the conclusions that follow from them" and to "relate the observations to other relevant studies" (ICMJE Recommendations 2019, 17).[23] Thus, in the original IMRAD schema we find two functionally specified sections framing two topic-dominated sections, which are functionally mainly descriptive.

As for the internal structure of the sections, the general principles of local organisation apply which I have dealt with earlier on. There are, however, preferred "moves" and also preferred sequences of these moves – and so-called "steps" within these moves – which scientists tend to make in presenting their results in a research paper. These aspects of the internal functional organisation of the IMRAD sections have been shown in research by Swales and his associates. For the introduction section, Swales described the following moves and steps:[24]

Move 1	**Establishing a territory**
Step 1	Claiming centrality
	and/or
Step 2	Making topic generalization(s)
	and/or
Step 3	Reviewing items of previous research
Move 2	**Establishing a niche**
Move 3	**Occupying the niche**

Thus, we see that the established practice of writing research articles reporting experimental research relies on fairly detailed models – including topic management – which a scientist can follow in presenting his work.

After these detailed observations on the IMRAD format it is, however, necessary to emphasise that reporting experimental research is not the only form of scientific or scholarly presentation of research. In many disciplines and subjects, we find papers which primarily present theoretical reflections or papers

[23] The ICMJE is the International Committee of Medical Journal Editors.

[24] See Swales (1990) and Swales (2004, ch. 7). I present the "steps" here only for Move 1.

that present empirical results which are not based on experimental work. This is, of course, true of fields like philosophy, history or wide areas of research in linguistics. A characteristic research article in Generative Grammar, for instance, might consist of the following elements:

(i) identifying a linguistic phenomenon (e.g., extraposition of adjuncts),
(ii) describing the generally accepted theoretical account of this phenomenon,
(iii) showing internal problems of this account,
(iv) presenting examples that are not covered by this account,
(v) presenting an alternative solution to the theoretical problem identified,
(vi) showing the advantages of this new solution for the sector of grammar involved.

A member of the Generative community will recognise this text as a globally coherent specimen of a research article in the tradition of their discipline. Thus, this example shows that scientific disciplines and "academic tribes" sometimes develop their own specific practice of writing research articles, including the practice of topic management. A young scholar being socialised into such a community, e.g. as a PhD student, will learn, among other things, the criteria of coherence involved in writing and judging research writings.[25]

5.6 Topic management in scientific writing

In the last section I dealt with criteria of coherence provided by text types, using as examples medical package leaflets, research papers of the IMRAD format, and a research paper in Generative Grammar. In the present section, I shall now continue with topic organisation in academic writing, however with an emphasis on individual strategies and measures of topic management in longer writings, mainly books.

5.6.1 Giving orientiation by making explicit aspects of topic management

In many contexts, authors of scholarly books consider it necessary or at least useful to make aspects of their topic management explicit by telling the reader what the topic of a particular passage is, where a given topic will be dealt with,

[25] For the concept of "academic tribe", see Becher/Trowler (2001). For features of disciplinary discourses, see Hyland (2004).

when a topic will be continued elsewhere, or that the topic was already dealt with earlier on in the book. I shall start by giving a few examples of utterances of this type, which could easily be multiplied:[26]

(1) The topic of this book is a sociology of philosophies […]. (Collins 1998, 12)

(2) The first three chapters present the general theory. (Collins 1998, 15)

(3) In the Epilogue I will argue at greater length that social construction of knowledge is realism, not anti-realism […]. (Collins 1998, 8)[27]

(4) IN THIS CHAPTER, I shall complete the exposition of Frege's doctrine concerning sense and reference; of the points here discussed in preliminary form, some will be taken up in greater detail later. (Dummett 1973, 152)

(5) This section deals with issues of information packaging. (Cambridge Grammar 2002, 1443)

(6) Let us turn, finally, to the conditional. (Dummett 1973, 338) [A subtopic of the logic of assertion, GF]

Whereas (1) to (3) give a general indication of the topic of a book and announce the topics of a series of chapters and a topic to be dealt with in the Epilogue, (4) to (6) are typical opening moves for chapters, sections or subsections, which orient the reader towards the immediate development of the topic. Numbers (7) and (8) are special in that they connect the topics of two adjacent sections:

(7) Let us turn now from these apparent exceptions to phrases of the form "most As". (Geach 1968, 89)

(8) Having looked at elementary discourse units and their identification, we now take the next step and consider the problem of recognizing coherence relations […]. (Stede 2012, 97)

A similar pattern, however with an extra indication of considerations of knowledge building is represented by the following quote from Goffman's "Forms of talk":

(9) Once we have considered the differential impact of system and ritual constraints upon talk, we can go on to consider a more complicated topic, namely, the inversionary effects of both these sets of constraints. (Goffman 1981, 26)

Another variant consists in announcing a subtopic and then announcing a brief passage that is inserted before beginning this new subtopic:

---

[26] Various types of "text-organizing metadiscourse" (in four different disciplines) are discussed in Hyland/Jiang (2020). Strategies of topic management in a late 19th century medical textbook are analysed in Gloning (2012).

[27] In this example the author announces both a functional aspect ("I will argue") and the topic of the Epilogue ("social construction of knowledge").

(10) Before turning to a discussion of the uniqueness implications that are central to the predictions of both Parsons' and Cooper's (as well as Evans') analyses, let me briefly comment on the major difference between them. (Heim 1982, 56)

(11) I want now to raise the issue of replies and responses but require a preface to do so. (Goffman 1981, 29)

Even more complex information on aspects of topic management is given in the following example (12):

(12) Before we continue our exploration of the age of genomic information with a look at the race to sequence the human genome, it is important to be more familiar with the relationship between genes and genomes. (DeSalle/Yudell 2020, 47)

This sentence occurs at the beginning of the second paragraph of a chapter with the heading "Sequencing the Genome". The authors here *refer back* to the preceding chapter containing an introduction to "the age of genomic information", they *announce* the topic of the current chapter ("the race to sequence the human genome") – which continues the earlier topic – and *motivate* the insertion of information on the topic of "the relationship between genes and genomes" ("it is important to be more familiar with").

A different kind of move consists in telling the reader that a topic that could be discussed at the present point in the text will be deferred and dealt with in a specified section later:

(13) Let us put this question aside for a moment. […]. (Collins 1998, 74f.)

(14) Questions about particular causal chains to one side for the moment, a second question concerns the *regularity* of both sorts of causal factors. (Hull 1988, 385)

(15) [...] a class of exceptions (about which I shall say more later) that is generated by the need to provide […]. (Mackie 2014, 40)

(16) This topic will be taken up again in in section 9 below. (Jackson 2020, 207)

Cross-references covering a longer distance in the text refer the reader to earlier passages, for instance in the case where a topic is readdressed that was already dealt with before, maybe more extensively. Sometimes such a hint is also meant to indicate to the reader that the writer is aware of the repetition of the topic.

(17) For instance, to return here to the topic I explored earlier, that of our connection to nature […]. (Somerville 2006, 181)

(18) Let us return here to the topic of the Oedipus complex itself. (Friedman/ Downey 2008, 112)

All these moves making explicit certain aspects of topic management, and others, provide orientation to the reader and contribute to the connectivity and smooth flow of the text.

5.6.2 Problems of global topic management and understanding

The moves mentioned just now are frequently solutions to problems of topic management. For instance, if, for whatever reason, it appears necessary to split the discussion of a topic into two parts, it might help the reader to be explicitly alerted to this fact. Unfortunately, writers do not always anticipate problems their readers might have with the topic structure of their writings. In such cases readers might fail to see the coherence of the text and therefore fail to gain a satisfactory understanding of the text.

The following example shows how a certain global topic organisation can create comprehension problems for the readers. My example is the book "Fallacies" by C.L. Hamblin, published in 1970, a classic of Informal Logic, which influenced many readers, including myself.[28] In this book Hamblin starts with a survey of the standard treatments of fallacies in the tradition of logic, followed by a long historical and systematic chapter on the Aristotelian tradition of this issue, a discussion of "arguments *ad*" (e.g., the *argumentum ad hominem*) in early modern logic, and a chapter on the Indian tradition of logic. This again is followed by a chapter on formal fallacies and a chapter on the concept of argument, where the systematic point of view takes over from the historical perspective prevalent up to this point. Now we have reached chapter 8, titled "Formal Dialectic", where Hamblin introduces a dialectical system of rules for argumentation dialogues. Finally, chapter 9 deals with problems of ambiguity in argumentation.

This brief listing of the topics dealt with in the course of the book already gives an idea why some readers might have found it difficult to find the guiding thread of the book. This (initial) experience lead Ralph Johnson, a prominent representative of Informal Logic, to write an article on "The coherence of Hamblin's *Fallacies*". Early on in his article, he writes:

> The question that occurs to me is: Why does the path that begins with a treatment of the "Standard Treatment" in Chapter 1 end up with a treatment of Equivocation in Chapter 9? The "logic", as it were, of this plan of development has not been made clear enough by Hamblin. The result is that the fundamental coherence of "Fallacies" is not as evident as it could be (Johnson 2011, 306).

[28] *fallacy* (in logic) 'a flaw that vitiates an argument' (Concise Oxford Dictionary).

What Johnson does in his article is to "reconstruct" Hamblin's topic management, thereby showing the fundamental coherence of the book. I shall not reproduce here Johnson's subtle reconstruction, but only hint at some of the results. To understand the topic structure of the book one has to realise that chapter 8 on "formal dialectic" is really the central chapter, to which all the others lead up. This chapter tries to show how an adequate theory of argumentation avoids the mistakes of the traditional treatments of fallacies. The historical chapters give a panorama of the kinds of problems to be dealt with in such a theory of argumentation, and the chapter on the concept of argument prepares Hamblin's own theory, which takes the form of a system of dialogue rules for arguments. Finally, the chapter on equivocation presents an *application* of his theory.

After this elucidation, we find that Hamblin's book can be represented as a coherent and inspiring text. The trouble seems to be that the author does not enlighten the reader at an early point in the book about the plan to be followed.[29] But this is not quite true. In fact, once one sees the basic structure of the book, one tends to notice small hints that could give the attentive reader a cue to the aim of the book, like the following:

> Work of this kind is a contribution to the theory of the use of language in practical situations: what Carnap calls Pragmatics and what we shall find reason to call Dialectic. It may be in this field that the discussions surrounding some of these so-called Fallacies find their true modern home (Hamblin 1970, 40).

As a general hint concerning the main topic of the book, this is, of course, quite understated and therefore, maybe, just not enough.[30]

It is remarkable that many influential texts should have exerted their influence in spite of a somewhat opaque topic management. This is true, for instance, of Austin's "How to do Things with Words", Wittgenstein's "Philosophical Investigations", and Grice's lectures on "Logic and Conversation".

An improvement of the topic management in Darwin's "The descent of man" (1871) was suggested by Richard Dawkins. He wrote: "It would have seemed natural to publish chapters 8 to 18 as *Selection in Relation to Sex* followed by a second book, *The Descent of Man*, consisting of the present chapters 1 to 8, and 19 to 21. That's a neat split into eleven chapters for each book, and many have wondered why he did not do this. […] The ostensible

---

[29] Alternatively, or in addition, the author might have placed the chapter on formal dialectic much earlier in the book, thereby indicating to the reader that this is the central chapter.

[30] The type of problem indicated just now is aggravated in connection with historical texts. For a "reconstruction" of the topic structure and coherence of an historical text, namely Leibniz's „New essays", see chapter 9.

connection between sexual selection and the descent of man is that Darwin believed the first was a key to understanding the second" (Dawkins 2003, 63). Therefore, it would have been helpful to present the key first.

The structural difficulty diagnosed regarding Hamblin's text is quite frequent in writings that combine a theoretical account of a problem with a historical perspective on the treatment of this problem. A presentation of the history of a problem usually presupposes some knowledge of the nature of the problem, which cannot be provided without some theoretical discussion. On the other hand, in some cases the theoretical discussion of the problem presupposes some knowledge of the requisite history. This is, for instance, frequently the case in philosophy. So the two topics – theory and history of the problem – seem to be entangled in a way that is difficult to unravel, mainly on account of problems of knowledge management.

As is normal with this type of difficulty, there are various different kinds of solution, which I cannot explore here in detail. I shall, however, give a few hints. One solution that could easily lead to problems of coherence would be to continually go back and forth between the two topics. Another strategy would be to start with the history, maybe giving a glimpse of the theoretical ramifications, and hope for the best. Or one could start with the theory, hint at the historical background, and leave the elaboration of the historical perspective for later. In any case, it would probably help the reader and contribute to the coherence of the text if the author would make explicit the difficulty in an introductory passage, introducing small information elements at relevant points, and generously using cross-references to connect the two perspectives.[31]

An interesting variant of this problem was noted by Quine in his *Word and Object* (1960):

> […] vagueness is an aid in coping with the linearity of discourse. An expositor finds that an understanding of some matter A is necessary preparation for an understanding of B, and yet that A cannot itself be expounded in correct detail without, conversely, noting certain exceptions and distinctions which require prior understanding of B. Vagueness, then, to the rescue. The expositor states A vaguely, proceeds to B, and afterwards touches up A, without ever having to call upon his reader to learn and unlearn any outright falsehood in the preliminary statement of A (Quine 1960, 127).

This is probably a familiar problem – and a useful solution – for every lecturer.

[31] In fact, I experienced a similar kind of problem in my treatment of coherence in monological text and dialogue, which I briefly reflected on in my introductory chapter.

5.6.3 Creating new topics

Thematic innovation is a special form of creativity. In science and scholarly work, discovering a new topic or a new combination of topics can lead to the development of a new branch of scientific exploration. An interesting example from the philosophy of language is Kripke's series of lectures on "Naming and necessity", presented in 1970. In this work, the title of which alludes to Carnap's "Meaning and neccessity" of 1947, Kripke creates a connection between the theory of proper names and the clarification of concepts like necessity and the a priori. He himself obviously assumed that for many of his colleagues this connection would appear new and maybe obscure. In the first sentence of the printed version of these lectures he therefore reflected on this problem with a touch of irony: "I hope that some people see some connection between the two topics in the title. If not, anyway, such connections will be developed in the course of these talks" (Kripke 1972, 253). This text triggered a lively debate, which continues to the present day, thereby establishing a new subtopic of the theory of reference.

Many scientific discoveries can be seen as partly consisting in the development of new topics of research and controversy, e.g., Lavoisier's discovery of a new theory of oxidation in the 1780s with its accompanying changes of terminology, which inspired a vigorous debate on this topic and a multitude of new developments in chemistry.

5.7 Topics: Some concluding remarks

In this chapter, I tried to show that topic structures and topic management are fundamental factors of discourse coherence and that, therefore, knowledge of topics and topic structures is a basic resource for producing and understanding coherent contributions to discourses. Topics are both a resource and an achievement. In many kinds of discourse, rules of topic management are fairly open and flexible and only rely on general principles of knowledge building and topic management, e.g., the principle that closely related subtopics of a topic should be placed in adjacent sections of discourse and that topic "leaps" should be explicitly indicated. On the other hand, for many formal text types fairly strict prescriptive rules of topic management have evolved and have been specified in regulations and guidelines. In many cases, principles of topic management are discipline-specific or community-specific, so that novices to such disciplines or communities have to learn these principles as part of their socialisation.

A final aspect worth emphasising at this point is that from the observations made in this chapter it follows that topic structures and topic management form important links between the *local* coherence of paragraphs, individual passages of texts, and chunks of dialogue on the one side and the *global* coherence of major parts of texts, whole texts, and whole conversations on the other. Thus, the claim made at the beginning of this chapter – that topics can be considered a fundamental organising principle of discourse – seems to be validated.

6. Text/image configurations: types of connectivity

6.1 Sequential uses of text and image

In my introductory section on text and image (2.6) I mentioned two types of the combination of text and image, namely sequential and simultaneous uses. These two types show two different kinds of connectivity. Whereas the sequential type may show linear coherence, very much like a text-text sequence, the simultaneous type shows a more complicated, nonlinear kind of connectivity, which I shall discuss later in this chapter.

I will start this section by repeating the following two examples of sequential text/image patterrns:

(i) A can describe an object and then *illustrate the description* by showing a picture of the object.
(ii) A can show a picture of an object and then *explain the picture* by drawing attention to features of the picture.

These are comparatively simple patterns that can, in principle, be described by the theoretical means introduced so far. As for the *coherence* of linear sequences performed according to these patterns, there remains of course the question which kind of picture can serve this purpose in the case of (i), and which kind of explanation is relevant to the picture in the case of (ii). In general terms one can say that in (i) the picture used for the illustration ought to show aspects of the object that are *relevant* for the purpose of the description. In the case of (ii), the explanation should refer to aspects of the picture that are relevant for the purpose of the act of showing the picture. To give an example of each of these patterns: If the purpose of the description is to draw attention to the beauty of a particular person, the picture following the description should show typical beauty features of this person. If the point of showing a particular picture, e.g., a portrait like the one shown below,[1] is to draw attention to the practice of using a landscape as background in Renaissance portrait painting, the explanation following the picture should concentrate on this topic. This kind of sequence is frequently found in scientific Powerpoint (and similar) presentations.

[1] Francesco del Cossa: Portrait of a man. (Public Domain, https://commons. wikimedia.org/w/index. php?curid=7734729).

Another point worth noting about these text/image sequences is the *knowledge* being built up by showing a picture or by describing an object. In the first place, this knowledge is a *resource* for the next action in such an action sequence. For instance, if certain aspects of an object are shown in the picture, these aspects can be referred to in the subsequent explanation. After looking at the portrait shown above, the viewer of the painting can be expected to know that the man holds a ring in his left hand and that there are mountains in the background. These things can now be referred to by expressions like *the ring held in his left hand* and *the mountains in the background*. On the other hand, this knowledge also *restricts* the kinds of statement the author can make after showing the picture. (He cannot, for instance, say without producing an inconsistency: *It is characteristic that there are no persons to be seen in the background*.) The same is true of the description presented before showing the picture. The details mentioned in the description may help the viewer to focus on them when they are shown the picture. On the other hand, if the picture does not "fit" the description, we also have a case of incoherence.

6.2 Interaction of simultaneous uses of text and image

In the following, I shall concentrate on the second variant of communicative action by means of text and image mentioned above, viz. the *simultaneous* use of text and image, which is represented, for instance, by the following pattern:

(iii) A can *introduce* a topic *by* naming the topic *and simultaneously* showing a picture symbolising the topic.

As opposed to the patterns mentioned in the last section, the combination of text and image is nonlinear, which presents particular tasks for the recipient. The basic question we are confronted with when dealing with such multimodal documents is the following: "How do recipients integrate the different modalities […] into a coherent meaning?" (Bucher 2017, 91). An answer to this question will at the same time elucidate the particular kind of connectivity to be found in such configurations.

The following is an example of this kind of text/image configuration from the "Headlines" section of "The Guardian" online (22 April 2021):

In this multimodal document the general topic is introduced by the noun phrase "Climate crisis", which is emphasised by the red colour and which also functions as a link to the detailed article. On the right-hand side of the document space, we see an industrial site emitting smoke (containing carbon dioxide etc.), which could, at first sight, be understood as representing a particular industrial site as the focus of the report. Once we read the lead sentence following the topic phrase, however, ("US vows to cut its emissions …") we will probably infer that the image is a stock photo showing a generic industrial site, an image that is iconic of the climate crisis. So we have to revise our first understanding of the picture, and we now take it to be an additional attention getter, generically indicating the topic of climate crisis. And the real topic of

the article turns out to be President Biden's "Climate summit", which took place online on April 22nd and 23rd, 2021.

The result of this rather elementary reconstruction of the reception process involved in reading the text/image configuration is that the reader starts by *selecting* a starting point, in my case the highlighted headline "Climate crisis", and from there goes on to look at the picture on the right-hand side, which they interpret as a documentary photo of a recent industrial problem site illustrating the topic of the text.[2] Returning their gaze to the lead section on the left, the reader arrives at a new interpretation of the photo and of the topic of the article in general. Thus, what we have here is a back-and-forth movement of the recipient between different textual elements, in which they develop an understanding of the document step-by-step. The path of this movement can be likened to sequences of the types "text T & illustration of T by picture P" and/or "picture P & explanation of P by text T", which shows a possible relation between our two types of connectivity.

The reception of multimodal documents has been studied systematically in reception laboratories, using different combinations of methods: eye tracking, thinking aloud during and after the reception process, re-narrations, knowledge tests, interviews, or questionnaires. In the following, I shall report some results of a research project by H. J. Bucher and his associates (Bucher 2017), which allow reconstructing the meaning-making process of the recipient, at the same time demonstrating some of the structures of multimodal discourse.

The example presented here is an analysis of eye-tracking data, which can serve as a window to reception processes. The text analysed is the front page of the German newspaper *die tageszeitung* of June 5, 2008, whose lead story covers Barack Obama's victory in the primaries in 2008 and his nomination as candidate for the presidency by the Democratic Party. The following figure shows a scan path of a recipient reading of this front page:

[2] It has frequently been assumed that recipients generally tend to look at the picture(s) first and start their reception from there. But this is not necessarily so, as reception studies have shown.

The title page shows a photo of the White House, and above this representation there is the headline "Onkel Baracks Hütte" (translated: "Uncle Barack's Cabin"). The lead text below the picture is as follows (my translation):

> Barack Obama makes history: Never before did a black politician have the chance to become the President of the United States. Obama secured for himself the nomination as candidate for the Democrats – in a duel with Hillary Clinton. Now he will attack the Republican John McCain in the battle for the White House.

The lines and bubbles represent the results of the eye-tracking procedure. The numbered bubbles show the „fixations" in their order and the lines represent the eye-movements from one fixation to the next.[3] What the zig-zag path of the eye-movement lines shows is that the recipient does not only perceive the different elements, headline, lead text, and picture, sequentially one after the other but also "*alternately* and *repeatedly*: The recipient attempts to get an

[3] Fixations are the time spans during which the eye remains still and processes the visual information of a specified area of interest. The size of a bubble indicates the length of the fixation.

overall meaning by taking each element as a context for the others" (Bucher 2017, 94). "With each step in the reception process, the meaning of the single element is expanded and, therefore, the context for the other elements is changed. Only due to this interpretive dynamic does the repeated fixation of the *same* element make sense" (Bucher 2017, 97). To give an example: The recipient seeks the meaning of the picture of the White House in this context by referring to the headline. And this again is elucidated by looking at the lead passage below. From there he/she then returns to the picture or the headline. Thus, this process is iterative and recursive. We can, however, assume that experienced recipients use certain routines like trying to find out the topic by looking at the headline – which in this particular case does not guarantee immediate success. Looking at the reception paths of the readers, we could say that the recipients convert the nonlinear document into a linear, or maybe more correctly, into a multilinear format.

What is important for comprehending this multimodal ensemble is also a certain amount of prior knowledge and of knowledge being built up in the process of the "interpretation". In order to understand the headline, the recipient has to know that "Onkel Toms Hütte" (German for "Uncle Tom's Cabin") is the title of a novel by Harriett Beecher Stowe, one of the most famous anti-slavery novels, and that Barack is the first name of the elected candidate. They must also recognise the photograph as a depiction of the White House in Washington. (Some of the recipients knew neither one nor the other.) Returning to the headline from the lead text and the photograph, and with the requisite knowledge, the recipient could *infer* that the headline is a humorous play with words, alluding to the title of the book and capitalising on the fact that the White House is no "cabin". This contrast also emphasises the historical dimension of the event. Thus, we have to assign the process of inference an important role in the interaction between the recipient and the co-deployed multimodal elements of a document. The inferential model of multimodal reception also demonstrates "that the meaning making of the author has to be distinguished from that of the recipient" (Bucher 2017, 97).

As for the theoretical status of this kind of inference, one could draw on Grice's theory of implicatures and his "pattern for the working out of a conversational implicature" (Grice 1989, 31). Assuming that the author of this multimodal combination did not seriously and truthfully want to assert that the White House is somebody's cabin, but nevertheless intended to make a relevant communicative contribution, and taking into account the information given in the lead text and the knowledge about the title and the content of the novel, one could conclude that the author wanted to playfully indicate the historical dimension of the step from living in a cabin to living in the White House. In addition, this implicature may have drawn on the practice of

ironically calling a luxurious abode a *Hütte* in German informal speech.

Returning to the author, we now have to ask: How can the author *trigger* the requisite process of interaction by deploying multimodal elements in a certain way? Obviously, it is not possible to anticipate all the different reception paths of different recipients that can be shown by eye-tracking studies. It is, however, possible to gain experience in orchestrating multimodal material in such a fashion that a certain configuration is a hopeful candidate for successful communication. Such experience is historically acquired by the continuous practice within communities of communicators. Although communication is always experimental to a certain degree, we find that certain preferred strategies and routines develop incrementally, for instance, in communities like those of scientists or journalists. In the sciences, for instance, the use of text/image configurations developed over the centuries. As early as the 16th century, we find subtle forms of text design with varied typography, diagrams and various kinds of illustrations, including, for instance, coloured images of plants (cf. Gloning 2015). A fairly recent development in science communication is the use of (Powerpoint) presentations, the development of which can be observed since the 1990s and which scientists had first to learn how to use, experimenting with various combinations of slides and (spoken) text.[4] As for newspapers, the amount of pictures used has "visibly" increased during the last 30 years, including documentary photos, infographics and sophisticated text design.[5] It is therefore not surprising that we find well-established patterns of the orchestration of different modal resources, especially text and images, in various text types, for instance newspaper articles, science presentations, and the city portraits and travel brochures mentioned earlier on.

Now that the observations made in the last sections on the use and reception of multimodal documents have hopefully answered some of the questions concerning the nature of the intermodal, nonlinear connectivity of text and image in such configurations, we can go on to an even more complicated type of configuration, namely the combination of sequential and simultaneous use of text and image in a multimodal narrative.

---

[4] For an eye-tracking analysis of Powerpoint presentations, see Bucher/Niemann (2012). Another recent development in science communication is the use of audiovisual material, e.g., on YouTube (see Boy/Bucher/Christ 2020; Bucher/Boy/Christ 2022).

[5] For examples of „text design" in newspapers, including various types of images, cf. Blum/Bucher (1998).

6.3 Connectivity in a multimodal narrative

A narrative of this type can follow the following pattern:

(iv) A can *tell a story by* describing a series of scenes *and simultaneously* showing a series of images representing scenes.

As an example of this complex type of pattern one could adduce the Bayeux Tapestry, which was produced, probably in 1070, after the army of William of Normandy had defeated and killed the English King Harold in the Battle of Hastings. The tapestry narrates the story of this battle and its antecedent history from the Norman point of view. The following image is a section of the tapestry that presents two scenes following after the death of King Edward (the Confessor) in January 1066, whose successor Harold was.[6]

In the first scene, Saxon noblemen (with weapons) present Harold with the royal crown, and then, in the second scene, Harold is seated on the throne with the royal insignia. To his left (on the picture) are two men showing their allegiance to the king and to his right (on the picture) is Archbishop Stigand, by whom Harold had been crowned king.

To understand this sequence of scenes one has to see that it follows a variant of a typical narrative sequencing pattern, namely the pattern "presenting the preparation of an event & presenting the event". In this case, the variant consists in the fact that the event as such, namely the coronation, is not shown, whereas its outcome, namely the situation of Harold sitting on the throne, *is* shown. What is also shown is an *aspect* of the coronation, namely the

[6] http://anglosaxon.archeurope.info/index.php?page=scene-37-2.

Archbishop Stigand, who performed the coronation. Understood in this way, the sequence of scenes provides the recipient with a coherent narrative sequence.

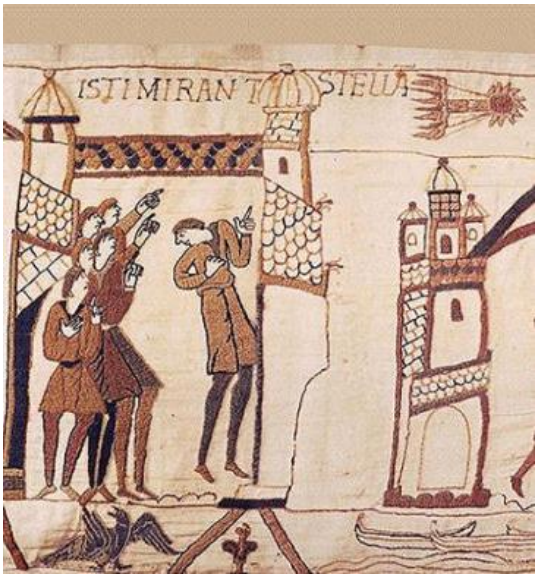
However, if we only had the sequence of pictures, we would have to know or to guess who the relevant personages are and what kind of situation or event is depicted. The group of men bearing arms on the left side could, for instance, be interpreted as a group of assassins preparing to kill the king. (This kind of interpretation problem is not unknown to the historian of art or the archaeologist.) In the case of the Bayeux Tapestry, we are in the fortunate position to have available not only pictures but also text. The Latin text explaining the scenes is inserted above the persons:

> (left) [hic] DEDERVNT: hAROLDO: CORONA[M] REGIS
> (Here they gave Harold the royal crown)
>
> (right) hIC RESIDET: hAROLD REX: ANGLORUM – STIGANT ARChIEP[iscopu]S
> (Here is seated (on his throne) Harold, King of the English – Archbishop Stigand)

An interesting aspect of this text is that by using the deictic pronoun "hic" ('here') the author of the text explicitly refers to the picture below. So what we have here is the second kind of connectivity, namely a connection between simultaneously presented different modes, in this case even explicitly indicated. (A similar kind of deictic procedure is frequently used in Powerpoint presentations, where the speaker uses a laserpointer to point to an aspect of an image, e.g., a part of a technical apparatus, indicating this aspect with an expression like *here we see* …) An important part of the verbal commentary on the tapestry is the identification of the persons involved, i.e., Harold and the Archbishop. Interestingly, the Saxon nobles are not identified as such, so the medieval recipient or the modern historian would have to know that the people presenting the crown are Saxon nobles.

As for the general meaning of these scenes, there is an interesting aspect still missing which only becomes clear if we know the whole story told by the tapestry. Seen from the Norman point of view, Harold's coronation was illegal, as he had earlier, i.e., before Edward's death, pledged fealty to William, Duke of Normandy, an event that is presented in two earlier scenes of the tapestry. In addition, Edward was said to have, on his deathbed, designated William to be his successor. This illegal character of the kingship of Harold is also indicated by one of the following scenes, where some men are presented looking and pointing towards a comet. As a comet, in the Middle Ages, was mostly considered a bad omen and a portent of disaster, this scene can be understood as foreshadowing Harold's defeat and death, a consequence of his illegally

seizing the throne.[7] The Latin inscription is: ISTI MIRANT(UR) STELLA(M) ‚These (men) wonder at the star'.



An interesting aspect of the coherence of this scene with earlier and later scenes is that without the knowledge of the interpretation of comets as portents of disaster and signs of God's wrath we could tend to see the comet scene as a strangely isolated scene without connections to the earlier and later story. However, *with* this knowledge the scene acquires an important function in the rhetoric of the whole story.

The take-away from these observations is that a picture can be used to show different things, depending on its context of use. In the case of the coronation scene, it can either show "a newly crowned king on his throne" or "an illegally crowned king on the throne". In the case of the comet scene, it can show an isolated historical event or an event with an important function for the story.[8] And it is only in the second interpretation that the scene is a coherent contribution to the story. As for the coherence of the total narrative, it is *the sequence of text/image configurations* that creates a coherent narrative.

---

7 http://anglosaxon.archeurope.info/index.php?page=scene-38-2

8 Halley's comet, which the scene probably refers to, appeared by the end of April 1066, the battle of Hastings took place in the middle of October 1066.

6.4 Act patterns for the use of text/image configurations

Let us now return to the second kind of connectivity, the connection between the simultaneously presented modes, text and image. To account for this kind of connectivity, we can assume several types of *act patterns combining the use of text and image*, of which some are similar to the ones to be found in the sequential presentation of text and image, for instance the following:

(i) verbally identifying a particular entity *and simultaneously* showing a picture of this entity

This is the pattern we find in newspapers or television news when persons are newly introduced, e.g., a newly elected politician or a so far unknown athlete having won a competition etc.

(ii) presenting the same topic by text and image, e.g., describing X *and simultaneously* showing a picture of X.

This is the type of coordination of text and image we found in the case of the mushroom (section 4.8.1 above), where both the description and the image are about the Penny Bun mushroom.

(iii) reporting an event *and simultaneously* presenting evidence for this event by showing a picture of the event.

This pattern is also frequently found in newspaper and television reports. What is interesting, and sometimes problematic, is the question whether the picture (or pictures in television) does (or do) in fact present evidence of the occurrence of the event or not. Frequently, in the case of pictures of uncertain provenance, it is quite uncertain what the picture really shows. In the case of an outbreak of violence, for instance, the picture might not clearly show who is attacking whom. In some cases, it may even be doubtful if the picture shows the relevant event at all and not some other event altogether. In addition, it is quite easy to manipulate images using digital application software.

As an unproblematic example I use here a report on the swearing-in ceremony of the US government in January 2021:[9]

[9] https://www.abc.net.au/news/2021-01-21/kamala-harris-sworn-in-as-us-vice-president/13076728. (21.01.2021)
A recipient may wonder who the persons wearing the black face masks are, and someone seeing this picture 20 years later might even ask what the meaning of these strange masks is. Are these persons members of an obscure cult? In 2021, of course, knowledge of the Covid pandemic was common knowledge.

Kamala Harris becomes first Black person, first Asian American sworn in as US Vice-President
Posted Wednesday 20 January 2021 at 9:12pm, updated Thursday 21 January 2021 at 12:57am

Kamala Harris becomes the first woman, first Black and first South Asian Vice President. *(AP/Saul Loeb)*

Basically the same type of pattern is also represented by the individual scenes of the Bayeux tapestry.[10]

These and similar typical patterns can be considered well-established configurations of text and image, and authors will use them routinely and recipients familiar with these patterns will recognise instances of such patterns straight away. And even if not, the routine pattern may suggest an inference about the meaning of a somewhat uncommon configuration of elements. In any case, the author's attempt at "recipient design" will include calculating the recipients' knowledge and also their powers of inference.

Closing this section, I want to summarise some aspects of the structure of the complex action patterns involved in the use of text/image configurations. The complexity of their structure can be made perspicuous by showing how the characteristic properties of complex actions conspire to create such a pattern. Typical patterns of this kind can be described by using the *by*-relation, the *and-simultaneously* relation, and the *and-then* relation mirroring sequentiality. The following simplified diagram shows how these structural properties are deployed in the case of a report on the event X:[11]

[10] A recipient may wonder who the persons wearing the black face masks are, and someone seeing this picture 20 years later might even ask what the meaning of these strange masks is. Are these persons members of an obscure cult? In 2021, of course, knowledge of the Covid pandemic was common knowledge.

[11] For a similar representation, see Bucher (2017, 111).

A informs B about X	*by* reporting how X happened	*by* using text T_1
	and simultaneously showing how and where X happened	*by* using a photo and a graphic representation
	and simultaneously indicating which elements belong to the report about X	*by* using a certain design element
	and then commenting on the reported event X	*by* using text T_2

6.5 Using text/image clusters

A special option of text/image design is the use of *clusters* of textual and pictorial elements, which is frequently found on the title pages of online newspapers and on start pages (or: home pages) of websites. As an example, I show here part of a title page of The New York Times:[12]

[12] https://www.nytimes.com/; Saturday, May 1, 2021. (01.05.2021).

The cluster consists of headlines (as links to the detailed articles), lead texts, and numerical infographics (on the development of global coronavirus cases and on U.S. and worldwide hot spots and U.S. vaccinations, as links to more detailed information). In the lower part of the cluster, which is mainly devoted to the situation in the U.S., there is also an additional interactive element, where the reader can search for the situation of the pandemic in their own county. The topic around which the different elements of the cluster are arranged is "global coronavirus cases", with two subtopics, namely the situation in India and in the U.S. The cluster design is a nonlinear arrangement, which relies on a segmentation of the total text and text/image material about a topic into different various "modules", which are (partially) arranged according to subtopics. In the upper two thirds of the above material, we find on the left a smaller unit about the new peak in coronavirus cases, consisting of a headline, a kind of short lead text, and the infographic vividly showing the peak. On the right-hand side there is a smaller cluster on the situation in India, with a commentary on Indian corona politics, a topical report about a hospital fire, and a report on the reactions of Indians living abroad. All these short elements are implemented as links to further information. This part of the opening screen could be termed a kind of teaser-only page. And on the lower third we find mainly infographics representing numbers of cases, hotspots and vaccinations, of which two are topographical graphics about the situation in different states of the U.S.

The advantages of this kind of design are generally assumed to be that it decomposes a complex topic into manageable portions, makes the complex topic surveyable, and allows the readers to select elements according to their respective interests.[13] The *connectivity* of such a multimodal ensemble of elements is mainly based on the topic structure and on patterns of the type "introducing a topic & elaborating on the topic". For the user, familiarity with the design used by "their" newspaper will facilitate orientation.

As I mentioned before, the cluster design in this kind of document is combined with a hypertext structure, i.e., many of the elements, e.g., the headlines and the graphics, are designed as links leading to other, more detailed elements of text. This kind of structure will be dealt with in the following chapter.

[13] The features of this kind of cluster design in newspapers are discussed in Blum/Bucher (1998).

7. Coherence in hypertext

7.1 Basic structures of hypertext

What a hypertext has in common with the image/text ensembles dealt with in the last sections is that it is a nonlinear text, or rather: a multilinear text. As coherence is typically defined for linear text, this seems to present us with a problem when asking in what sense we can speak of the coherence of hypertext. The following sections will try to give an answer to this question. Foreshadowing the results of this exploration, one can say that on closer inspection it emerges that some of the fundamental structural problems with hypertexts *are* classical problems of coherence.[1] To clear the way, I shall now introduce some basic concepts of hypertext theory.

A hypertext is a configuration of textual elements, including pictorial elements, that are connected by so-called links. Typical structures of such configurations are network structures or tree-like structures. A simple example of a hypertext network structure can be represented as follows.[2]

Let us assume we want to inform somebody about COVID-19. In this case we could write 7 textual elements (paragraphs), [T_1] to [T_7], each dealing with a relevant (sub)topic, e.g. the following:[3]

[T_1] (the disease COVID-19**)**
[T_2] (the virus SARS-CoV-2)
[T_3] (transmission of the virus)
[T_4] (therapies)
[T_5] (preventive measures)
[T_6] (vaccines)
[T_7] (testing)

Now we arrange these paragraphs in form of a network, with *nodes* representing textual elements and *edges* representing (potential) transitions from one element to the other (e.g. from [T_2] to [T_3]), assuming a connection between

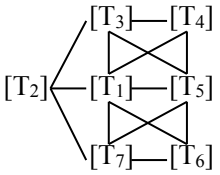
---

1 For earlier discussions of coherence in hypertext from an action-theoretic point of view, see Bucher (1999); (2001), Fritz (1999).

2 Tree-like structures are found, for instance, in hypertext narratives where there is a linear path of documents arranged chronologically plus various links scattered over the individual documents. An example of this kind of structure is Jeff Pack's autobiographical "Growing up digerate", where the series of computers in his biography forms a linear sequence. (https://cyberartsweb.org/cpace/infotech/digerate/TITLE.HTML). (11.02.2022).

3 A more detailed discussion of the structure of such a topic was given in chapter 5.

the respective topics of these elements. These edges of the network graphs we call the *links* of the hypertext. This gives us a network of the following type:



We further assume that we can move from one linked topic to the next in both directions. (This gives us an undirected graph.) Now the reader/user can follow the links from one textual element to the next and move through the network (of topics), e.g., from [T_1] to [T_2], [T_3], and [T_5], thereby producing a *sequence* of textual elements.[4] Such a sequence, together with the links connecting them, we call a *path* through the hypertext. As one can easily see, such a path resembles a linear text and can, therefore, be evaluated for its coherence properties. This is related to what has sometimes been called the "virtual textuality" of hypertext. Hypertext is *multilinear* text because there are different possible paths through the network. So here we have a first answer to the question from which we started this section: If we want to explore the coherence of hypertexts, we have to start by analysing the *coherence of reader/user paths*. What is different from normal linear text is that the user has the *choice* of following one path or the other. This is, of course, not *totally* different from the practice of reading books; we can, for instance, start in the middle of an academic book and jump from there to a later chapter and from there to the introduction etc. This may, however, be quite risky, as the author will mostly have organized the book in such a way that the reader profits by reading it sequentially.

What is still missing in this picture of hypertexts is the fact that the textual elements of such a configuration are presented in a digital format and that the configuration itself is electronically implemented, like, for instance, in the Internet. As for the links, they can be activated (in the simplest form) by clicking on a bit of text or an image (the "anchor"), which brings us from the *source* text/image to the *target* text/image. In the following extract from the Wikipedia article on COVID-19 the links leading to articles on the virus, the symptoms and the transmission are marked in blue (underlined in the version presented here):[5]

---

4 An important function of modern browsers is the *return* button, which permits the users to retrace their steps.

5 https://en.wikipedia.org/wiki/COVID-19.

(1) **Coronavirus disease 2019** (**COVID-19**), also known as **the coronavirus**, **COVID** or **Covid**, is a contagious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). [...] Symptoms of COVID-19 are variable, but often include fever, cough, headache, fatigue, breathing difficulties, and loss of smell and taste. [...] Transmission of COVID-19 occurs mainly when an infected person is in close contact with another person. [...].

What is remarkable here is that Wikipedia provides two functions of the link. When pointing the cursor on the link or hovering over the link a pop-up window opens, presenting the user with a brief description/explanation (e.g., of the symptoms of COVID-19) as a preview of the target article, whereas when clicking on the link the user is lead to the complete target article (on symptoms).

7.2 Links, paths, and local coherence

At this point we can ask another important question to guide our further exploration: How do users of a hypertext make sense of the path they are following? From what we have seen so far in this book the answer to this question must lie in the following direction: Users make sense of a path or a segment of a path by using their knowledge of principles of textual organisation, i.e., by regarding it as a realisation of a sequencing pattern and/or a characteristic topic structure, and by drawing inferences on the basis of their locally generated and generally available knowledge.

This leads to two questions: In which way are the source text and the target text related and in which way does the link indicate such a relation.

Let us start from the last example, where the paragraphs on Covid-19 form the source text. Here we have links in the form of marked noun phrases like *transmission of COVID-19*, so called "hot text". In this case the link expression indicates a topic which can be considered a subtopic of the general Covid-19 topic. So what can guide the user here is his knowledge of the topic structure of the Covid-19 topic or, more generally, of the structure of the topic *infectious disease*. *Functionally*, the description of the transmission in the target text follows the pattern "elaboration": *By* presenting passages describing transmission – in this sequence – the author elaborates on the topic of the disease. Now here we have an interesting point. If we reach the transmission text from a different topic in the network, for instance from [T_5] (*preventive measures*), we would ascribe it a different function: In this case it could contribute to *explaining* why certain preventive measures (e.g. social distancing) are efficient. This difference in function – elaborating vs. explaining – is not explicitly indicated, so

the user has to anticipate the respective function by inferring it on the basis of the source text and the topic indicated by the link expression. This is a general problem with this kind of link on the internet and may lead to comprehension problems.

As for the use of a pop-up window mentioned before, its function of meaning explanation or brief encyclopedic information is also performed by online dictionaries and glossaries and is frequently used on help sites or in online books, where individual words are linked to a dictionary. If, for instance, the unsuspecting reader comes across the word *sesquipedalian* they can just click on the word to be enlightened. In this case, the function of the link is easily recognised and anticipated. In other cases, this is less so. In any case, "[…] users *expect* purposeful, important relationships between linked materials" (Landow 1991, 83).

This leads us to the question of how different types of links and link labels can contribute to the comprehension and usability of user paths in hypertext. The question of link labels has been the target of lively discussion under the label of "link typing" (cf. Kopak 1999, Landow 2006). From the point of view pursued in the present book, links should contribute to coherence and to the easy flow of information by connecting textual and pictorial elements that are functionally and/or thematically related. In addition, the link label or link text can contribute to making explicit the textual organisation of a user path, including the functional organisation, the topic structure, and forms of knowledge building.

A general aspect of the use of labeled links is what Landow called the rhetoric of departure and the rhetoric of arrival or destination (Landow 2006, 14f.). From the point of view of usability, Nielsen gave the following advice (Nielsen 2000, 66):

> The rhetoric of departure. Clearly explain to users why they should leave their current context and what value they will get at the other end of the link.
>
> The rhetoric of arrival. Clearly have the arrival page situate users in the new context and provide them with value relative to their point of origin.

What is not surprising is that the types of links indicated depend on the text type or medium of the respective textual elements. I shall give a few examples here.

In newspapers, we frequently find a combination of two types of link leading to a detailed report article, namely a general topic-indicating label and a link labeled with the lead sentence or "teaser", which leads to an elaboration (extended treatment) of the topic in form of the detailed report. The following is an example of this kind of double link from The Guardian (09.05.21):

(2) **Colonialism** Church of England to consider removing or altering slavery monuments

A similar practice can be found on the front page of the New York Times shown above. Thus, readers familiar with this practice will not find it difficult to see the connection between the topic indicated by the topic-indicating label (and the lead) and the topic of the detailed report. At the same time, they will also understand the function of the link, namely to connect to an elaboration of the brief preview. If we return back from the extended report to the lead sentence by using the return button of the browser, the link represents the inverse relation, i.e., the link leads from the extended text to a short version, the lead, which in this sequence can be read as a kind of summary. But, of course, this functional relationship is not explicitly indicated in this case.

A related, but less specific kind of label is found in all kinds of texts where short versions of a textual element can be extended to a longer version, maybe for users with special interests, with a link labeled *read more* or *show more*, e.g., in the etymology section of the OED. As a rule, the user will interpret this link as indicating that they can expect more information about the same or a closely related topic.

Another interesting kind of link is the one used for giving *background information*. For example, in a report on the hijacking of a Ryanair flight, BBC News presented the following news text:[6]

(3) Several European airlines have said that they will not fly over Belarus, days after a dissident journalist was arrested on a flight diverted to Minsk. [...] Western countries accuse Belarus of hijacking the Ryanair plane carrying journalist Roman Protasevich on Sunday.

At the end of the paragraph there are three bullet points with the following links, printed in bold face, indicating background material, including information on the journalist:

- Belarus plane: What we know and what we don't
- Who is Roman Protasevich?
- 'Perfect storm' prompts EU to act fast on Belarus

This form of linking was introduced in the BBC News in 2008.[7] In the case of the information given about the journalist, this way of linking is an alternative to using the name of the person as an anchor for this information.

---

[6] https://www.bbc.com/news/world-europe-57239162 (25.05.2021).

[7] Cf. https://www.bbc.co.uk/blogs/theeditors/2008/08/new_ways_of_linking.html; the expression *perfect storm* was used by a European diplomat to characterise the reaction of European politicians.

Topic-indicating labels for links are frequent in online academic writing, where, for instance, the entries in the contents section of books are used as links to the respective chapters and sections. The same is true of online encyclopedias like the Stanford Encyclopedia of Philosophy. Scientific journals frequently provide links from in-text citations to the reference section as well as links to supplementary data, figures etc.

Generally speaking, typed links indicating topics are used much more often than those indicating functions. Landow mentions "specific kinds of relationships" like "exemplifies", "influences", "contrary argument", "derives from" as worth being explicitly indicated (Landow 2006, 18). These are however rarely explicitly named as link labels. This means that the users are frequently left to their own devices to infer the functional relation between two linked documents, even if a topic-indicating label is provided. In fact, users will frequently perceive only vague functional relationships between documents, e.g., addition, continuation, or specification.

A useful type of function-indicating link label is the question type label, which is based on a question-answer dialogue sequence and is frequently used on help and support sites. The following are examples from the Help page of the Manchester University Library:[8]

(4) How do I search for items in the Library catalogue?

(5) Why can't I access an e-journal, e-book or database?

These quite intuitive links lead to instructions and tips for using the library search facilities (4) and solutions to a problem (5). This type of link is also found in lists of "frequently asked questions". Another link label indicating a problem-solving textual element consists in the description of a problem, like the following:[9]

(6) You click the download button and nothing happens

Finally, there are also activity-indicating link labels like a button with the inscription

(7) report a problem

Links to explanations sometimes take the following form: *for an explanation click here*. Other types of function-indicating link labels include text-type indicating labels like *explainer* or *opinion*, which are used in newspapers.

[8] https://manchester-uk.libanswers.com/. (02.10.2021).

[9] https://support.jstor.org/hc/en-us/articles/115004798167-Troubleshooting-Downloading-a-PDF#clicked-download. (11.02.2022).

7.3 Navigation and global coherence

Whereas the links mentioned so far contribute to the organisation of the local structure and therefore (primarily) to local coherence, there is also a level of global hypertext organisation, to which I shall turn now. For the users, there are basically two tasks they have to solve: to successfully navigate the hypertext and to keep their orientation in hyperspace. For both these tasks hypertexts have to provide the user with means for solving them. Following Nielsen (2000, 188), the three fundamental questions of navigation are:

(i) Where am I?
(ii) Where have I been?
(iii) Where can I go?

One of the answers to the question "Where can I go?" is, of course, provided by individual typed links, as we have seen in the last paragraphs. This feature, however, permits the user to look just one step ahead. On a more general level, choices of where to go can be given by global navigation bars, either as horizontal "top bars", as vertical sidebars, typically positioned on the left side of the page, or as navigation boxes at the bottom of the page. Examples of horizontal navigation bars are the menus of "fields" given in newspapers like The Guardian: *News – Opinion – Sport – Culture – Lifestyle*. These high-level categories usually appear on every page of a website and give the user an idea of what is generally available on the site. A related function is served by tables of contents, which allow navigation within a document by providing an overview of the topics dealt with, e.g., in a Wikipedia article or an electronic book. An interesting version of a device providing an overview is the system of navigation "tiles" used on "The Victorian Web", a website dedicated to various aspects of Victorian literature and culture in general.[10] As an example I show, on the following page, the use of this kind of structure in the function of a general index on the homepage.

[10] https://victorianweb.org/index.html. (11.02.2022). Editor-in-chief and Webmaster is George P. Landow.

Another navigation feature of this website is the trail of "breadcrumbs" at the top of each webpage, explained as follows:[11]

> The furthest 'breadcrumb' on the left is a link to the homepage. But you may wish to try the intermediary ones before going there. The furthest right is to the section to which the current page belongs. In between are links to the wider context of your subject. For example, if you are reading about Dickens's A Christmas Carol, the breadcrumbs will look like this:
> [Victorian Web Home —> Authors —> Charles Dickens —> Works —> The Christmas Books].

In this case the sequence of topics mirrors the page hierarchy of the website, which is generally a frequent use of breadcrumb navigation. It allows the users to identify the page on which they currently are and to jump directly to higher levels of the hierarchy.[12] A second type of breadcrumb navigation are path breadcrumbs, which show the steps in the path the user has taken so far. This kind of navigation allows the users to answer the question "Where have I been?", thereby possibly reconstructing the coherence of their user path, and to retrace their steps, without using the "back" button.

Thus, navigation systems contribute not only to the usability of a website but also to the understanding of the global coherence of a website. This is, of course, particularly relevant in the case of complex, hierarchically organised websites like the Victorian Web mentioned before. Looking at a breadcrumbs sequence, for instance, the users may recognise parts of the global topic

---

11 https://victorianweb.org/misc/nav.html. (11.02.2022).

12 The user can, of course, also identify the current page by looking at the URL bar, in the Dickens case: https://victorianweb.org/authors/dickens/xmas/index.html.

structure and the functional steps that lead to their current position in the website.

An interesting special feature of hypertext concerns the way *knowledge* is built up in the path of a hypertext. As I discussed in chapter 3, in the case of linear text systematic knowledge building is the rule, i.e., the author tends to present information incrementally, keeping track of the knowledge provided so far and the knowledge necessary for introducing new information. If an author fails to comply with this principle, a text with knowledge gaps will be considered incoherent and at least partially incomprehensible. This is fundamentally different in hypertext. Hypertext documents connected by links basically have to be informationally self-contained. This is a consequence of the fact that a given textual element in a hypertext may be accessed from different directions, e.g., also from "outside" the respective website, which means that users taking different paths und therefore having different "histories" of knowledge acquisition should all be able to understand the respective textual element. For instance, let us assume an author writing a hypertext on genomics wants to include a paragraph on nucleic acid sequencing which should be accessible from different directions. In this case, they should either explain early on in this paragraph what nucleic acids are and what sequencing is or provide links within the paragraph explaining these two items. The author cannot expect the user to have learnt the meaning of these expressions somewhere earlier in the text.

This feature of knowledge building restricts the options of coreference in hypertext. Pronominal coreference of the usual type as well as the coreferential use of deictic expressions like *this king* or contextually bound noun phrases like *the king* are only possible *within* a hypertext document, especially within a paragraph. This is due to the fact that the anchor for a backwards-looking coreference may not be available. For reference to the same entity across documents other means are necessary, e.g., the use of a proper name or a definite description with fixed reference (e.g., *the Internet*). If in doubt whether the user will recognise which entity is referred to, the author can provide identifying knowledge "on the spot" by supplying an explanatory link. To give an example: in an online article on the Battle of Maserfield, the battle is introduced by the following sentence:[13]

(8) The **Battle of Maserfield**, was fought on 5 August 641 or 642, between the Anglo-Saxon kings Oswald of Northumbria and Penda of Mercia, ending in Oswald's defeat, death, and dismemberment.

[13] https://en.wikipedia.org/wiki/Battle_of_Maserfield.

Within the article, especially within paragraphs, it is possible to refer to the battle by using expressions like *the battle* or *this battle*. However, in a related article on King Oswald the complete expression *The Battle of Maserfield* has to be used, accompanied by a link explaining this term.[14] Thus, the accepted remedy for the lack of typical means of coreference is the provision of links supplying identifying knowledge.

7.4 Ten points about the coherence of hypertext

To sum up the results of my observations on the coherence of hypertexts:

(i) Hypertexts are multilinear texts. Coherence is primarily a property of linear paths through the hypertext.
(ii) Paths are constructed by means of links between documents/textual elements, including images.
(iii) Authors base the organisation of path options on organising principles of discourse, e.g., topic continuity/topic relatedness, functional sequencing and knowledge building.
(iv) The same is true of users following and tracking paths and making inferences in order to close gaps of understanding.
(v) The links between documents can be made explicit by *typing* the links, i.e., by indicating the topic or function of the target document or the functional/thematic relation between the documents.
(vi) Frequently, the relations between individual documents in a path are only vaguely defined. Within the field of topic-connectedness we find a gradient from strict thematic coherence to fairly loose thematic connections. Generally speaking, in hypertext we get everything from strong prototypical coherence in tightly organised websites to minimal coherence, for instance when aimlessly browsing.
(vii) In many document types there seems to be a preference for topic-indicating link labels as compared to function-indicating ones.
(viii) Knowledge building requires extra care in hypertext. Moving across documents by links, the author has to make the individual document informationally self-contained, e.g., by providing links to background information.
(ix) Usability and global coherence of a hypertext can be ensured by means of navigation devices that facilitate user orientation and comprehensibility of paths.
(x) In an extended sense one could call a hypertext coherent if it provides the user with functionally and thematically coherent paths.

[14] https://en.wikipedia.org/wiki/Oswald_of_Northumbria.

8. Communication principles, constraints, and strategies

8.1 A family of principles

In my introduction I suggested that a useful way of looking at coherence could consist in conceptualising it as the content of a communication principle (or maxim) that can be implemented in various forms and various degrees by using a family of textual resources.

At the same time, we can see that this principle is related to a whole family of other principles which can be thought of as either concretising the coherence principle or complementing it in its function as a standard for the quality of discourse. Some of these principles are fairly general, as for instance the principle of relevance, which I already briefly mentioned, others seem to be specialised to certain genres and domains of discourse, e.g., the burden-of-proof rule for (certain) controversies.[1] In addition it is worth noting that some of these principles seem to be historically variable. Take, for instance, certain politeness principles and items like the principle of tolerance, which seem to have grown in importance in the course of the 18th century (in Europe). The following short list is meant to give a first impression of what kinds of principles we have to think of in this context:

(i) the principle of connectedness
(ii) the principle of continuity
(iii) the principle of orderliness
(iv) the principle of relevance
(v) the principle of consistency
(vi) the principle of non-repetition
(vii) the principle of comprehensibility
(viii) the principle of recipient design
(ix) the principle of clarity/perspicuity
(x) the principle of explicitness
(xi) the principle of precision
(xii) the principle of informativity

Of these, the first five seem to be particularly closely related to the principle

---

[1] Together with other principles, this traditional maxim is listed among the "rules for critical discussion" in van Eemeren et al. (2002, 182f.). For the role of communication principles in (Early Modern) controversies, see Fritz (2008).

of coherence, whereas the second group includes principles related to comprehensibility and other aspects of good discourse.[2]

8.2 Communication principles: quality criteria and heuristics of interpretation

Readers of Grice's "Logic and conversation" (Grice 1989) will have noticed that in the context of my discussion so far, the communication principles/maxims have a different function from the one assigned to them in the context of Grice's theory. Whereas I have so far treated them as principles guiding speaker/writers in their text production, Grice uses communication principles as part of a heuristic of interpretation.[3] In Dascal's words this heuristic provides "a set of (at least partially) ordered rules that would provide the hearer (and the speaker) with the instruments for *guessing* (rather than *deducing*) the implicatures, but for guessing in an 'educated' or 'systematic', rather than haphazard way (Dascal 1979, 167). If, for instance, a contribution to a dialogue ostensibly violates the principle of relevance, the hearer could assume that in fact the speaker is implying something different from what they are explicitly saying, i.e., making a so-called *conversational implicature*. This assumption then "may trigger a process of looking for an alternative interpretation of the utterance […] which shows how, in spite of the appearance to the contrary, the utterance does in fact meet the conversational demand" (Dascal 2003, 628). To give an example from Grice:

> Suppose that A and B are talking about a mutual friend, C, who is now working in a bank. A asks B how C is getting on in his job, and B replies, *Oh quite well, I think; he likes his colleagues and he hasn't been to prison yet*. At this point, A might well inquire what B was implying, what he was suggesting, or even what he meant by saying that C had not been to prison. The answer might be any one of such things as that C is the kind of person likely to yield to the temptation provided by his occupation [etc.] (Grice 1989, 24).

Thus, the seemingly irrelevant remark triggers the search for what A "really meant" (i.e. his implicature). Part of the background of this procedure is the hearer's assumption that the speaker is following the principle of relevance. And by "solving" the riddle presented by the ostensible violation the hearer "saves" the coherence of the "question/answer" dialogue. This is the main

---

[2] For a more detailed discussion of various dimensions of discourse quality and the granularity of communication principles, see Fritz (2017, ch. 5).

[3] For this view of the role of Grice's principles/maxims, see Dascal (1979, 167); Levinson (2000, 27ff.).

aspect under which Grice's theory of implicatures is relevant to problems of coherence.

As I mentioned before, one can also – or even primarily – consider these principles as guiding principles for discourse quality. From my point of view this guiding function for speakers and writers is primary and the heuristic function is derived. Because the hearers/readers assume that the speakers/writers follow the principles, they can use apparent violations as an indicator of a possible "special" speaker meaning. An opposing view is taken by Levinson: "Instead of thinking about them as rules (or rules of thumb) or behavioral norms, it is useful to think of them as primarily *inferential heuristics* which then motivate the behavioral norms" (Levinson 2000, 35). For my particular purpose here, this difference plays no important role, so I will not discuss this problem further. The following observations pertain mainly to the function of communication principles as guiding principles.

8.3 The relevance of communication principles to aspects of coherence

Returning to my short list of principles given at the beginning of this chapter, I will now discuss some of the interrelations between these principles and the principle of coherence.

The principle of *connectedness* emphasises a basic aspect of coherence, namely that the individual utterances should be connected to one another in certain ways. One basic type of connection is the connection of adjacent utterances by grammatical and lexical means. However, the mere fact of connectedness in this sense does not guarantee coherence, as the elementary units of a text may be formally connected by connectors and referential means and still not be coherent in a strict sense, because they cannot be associated with sequencing patterns and/or lack an overarching topic structure. If we accept this view, then formal connectedness is a weaker sub-principle of the principle of coherence. If we extend connectedness to the use of all the relevant discourse resources (sequencing patterns, topics etc.), it is more or less equivalent with coherence. Various types of connectedness are discussed by psychologists, e.g. conceptual connectedness. Under this aspect a text is "fully connected if each proposition is conceptually connected to one or more propositions. In this case the text would satisfy a minimal formal criterion for coherence. […] A text with sparse connections is predicted to be difficult for a reader to understand" (Graesser/Britton 1996, 344).[4]

[4] See also Graesser et al. (1997, 179f.).

Continuity contributes to coherence in various ways. Maybe the most important is its role as default mode in topic management. This is true, for instance of narratives, where a series of reported events is held together by a continuous topic, e.g., the "point" of the story. In newspapers, a cluster of texts, including reports, statistics, commentaries, and letters to the editor, can be unified by their common topic. In conversations, a certain amount of topic continuity is also normal, albeit frequently combined with a strategy of small topic shifts. Another important type of continuity is continuity of reference to a salient entity, for instance in genres like biographies and obituaries, where continuous reference to the protagonist is constitutive, or in scientific papers, where a salient concept is introduced and continuously referred to in a series of paragraphs.

The principle of orderliness can be applied to various aspects of discourse.[5] Concerning the order of topics it may prescribe, for instance, that one should not jump back and forth between different topics. In this sense it is a variant of the principle of continuity. As for speech act sequences, it calls for continuous sequences, making discontinuous sequences dispreferred. If, for example, an argument for a claim is given, this argument should be placed in the immediate neighbourhood of the claim and not somewhere else in the discourse. However, there are textual devices that allow the production of discontinuous sequences, so this principle can be overridden. In complex texts, clarification moves can offset problems of orderliness. To give an example: In his paper on indicative conditionals, Grice introduces such a device in announcing and presenting a summary:

> It might help to clarify the somewhat tortuous course of my treatment of indicative conditionals if I specify, in a summary form, the sequence of ideas which have been canvassed in that treatment (Grice 1989, 83).

In psycholinguistic experiments the sentences of a paragraph were sometimes jumbled up to explore how this influenced the comprehension of the text. Although this is a rather problematic method, especially since the jumbled-up text is not "the same text" as its original, it shows – somewhat unsurprisingly – that the order of sentences is an elementary aspect of coherence.

As for the principle of *relevance*, I already mentioned earlier that this principle has a wider application than the principle of coherence. A connection between these two principles consists in the fact that in a coherent series of utterances these utterances are also relevant to each other. If a sequence of

---

[5] In the Gricean account, the somewhat vague maxim "Be orderly" is mentioned in his list of maxims under the heading of the supermaxim "Be perspicuous" (Grice 1989, 27), a maxim that goes back to classical rhetoric: "Nobis prima sit virtus perspicuitas" (Quintilian 1995, vol. 2, 149).

utterances S1 & S2 & S3 is coherent, S2 and S3 are relevant as succeeding utterances to S1 and S2 respectively, and S1 is relevant as a unit preceeding S2 (and probably) S3. Thus, coherence is a sufficient condition for (a certain type of) relevance. On the other hand, there are various types of relevance that are independent of coherence. A certain type of knowledge may be relevant to a certain speech act (e.g., identifying knowledge for acts of reference) without being in any sense directly involved in matters of sequential coherence. Similarly, a certain topic (dealt with in a different book) may be relevant to a given text without being in a coherence relation with speech acts of this text. Or: Two propositions may be relevant to each other in the sense that one proposition follows from the other. But if they are not sequentially connected, they are not coherent in the sense used in my account.

A particularly interesting principle is the *principle of consistency/ inconsistency*, which is frequently mentioned in the context of discussions of coherence. Logicians tend to emphasise that of two inconsistent (contradictory) assertions not both can be true (e.g. Geach 1976, 6f.). But an even more fundamental reason for avoiding inconsistencies (self-contradictions) is mentioned by Strawson in the first chapter of his "Introduction to logical theory": "The point is that the *standard* purpose of speech, the intention to communicate something is frustrated by self-contradiction" (Strawson 1952, 3). This is probably the reason why inconsistency is frequently considered a form of incoherence. However, this issue bears closer examination. The main difference between coherence and consistency lies in the feature of sequential order, e.g., in the case of two adjacent utterances, which is characteristic of coherence but not necessary for the (in)consistency of two speech acts. Inconsistent speech acts in direct sequential neighbourship seem to be quite rare, anyway. This is true, for instance, of two adjacent head-on self-contradictory assertions. The sequence "assertion$_A$ that p & assertion$_A$ that not p", if it occurs, must be considered a violation of a sequencing pattern for assertions. Sentences like *I like it and I don't like it*, however, are normally not used to make two contradictory assertions in a sequence, but to indicate that the speaker is in two minds about his attitude to something. A slightly different type of inconsistent sequence follows the pattern "assertion$_A$ that p & assertion$_A$ that A does not believe that p". This type of inconsistency, which I already mentioned in chapter 3, has been known as Moore's paradox.[6] If a person asserts that it is raining and in the same breath asserts that they do not believe that it is raining, this sequence is inconsistent in a different way from the explicit self-contradiction. In our framework the inconsistency and the feeling of absurdity can be explained as

[6] See Wittgenstein, "Philosophical Investigations", part II, section X. Wittgenstein gives a different explanation from the one I suggest.

follows: In asserting that p the speaker incurs the commitment that they believe that p. If they now add, as an adjacent utterance, the assertion that they do not believe that p, there is a contradiction between a commitment of the first assertion, which is part of the updated common ground after the assertion, and the second assertion. Thus, this is again a violation of a (general) sequencing rule and therefore a case of incoherence. The generalised principle could be formulated as follows: An utterance should not be inconsistent with commitments in the common commitment store. The interest of this example lies in the fact that it can be considered one of the motivations for the assumption of a commitment store in a theory of discourse.

So far, these examples are rather artificial, but there are other problems of inconsistency in discourse that are worth examining, mainly what one could call "distant inconsistencies". These are much more frequent and sometimes also more difficult to detect than the adjacent type.[7] If a hearer/reader diagnoses such an inconsistency, they can come to one of the following conclusions – and maybe others: (i) The speaker/writer changed their mind, (ii) they forgot what they said earlier on, (iii) they have some kind of cognitive problem, (iv) the hearer/reader must have misunderstood one of the two relevant utterances. The last conclusion, where it is relevant, is the most interesting, because it may trigger a query or a process of re-interpretation: "… what triggers a search for an alternative pragmatic interpretation of an utterance is the detection of a "mismatch" between its ostensive utterance meaning and the context" (Dascal 2003, 639). Another point worth noting is that in many cases inconsistencies are rather "soft" so that they can somehow be accommodated or are experienced as mere "tensions". And, of course, what may look self-contradictory to one person might be perfectly acceptable to another.

The principle of *non-repetition* is related to the principle of informativeness: Repeating a proposition that is already in the commitment store fails to advance the discourse and should therefore be avoided. This is, however, not the whole truth. In fact, "no-one is quite sure whether repetition is good or bad, either in literature or linguistics" (Aitchison 1994, 17). This can be demonstrated by pointing out various uses of repetition in texts that are useful and commonly applied. One such use is the repetition of a claim with the intention of emphasising the claim or insisting on it. Another function is the use of repetition as a kind of *leitmotif* technique. And in linguistics texts examples are sometimes repeated "for convenience of the reader". In all these functions and

[7] Another type of inconsistency that might not be easily detected is the case of inconsistency with remote consequences of commitments (see Hamblin 1970, 264). The problem of "contradiction detection" has been a research topic in both psychology and natural language processing (see, for instance, Kintsch 1998, 181ff.; Marneffe et al. 2008).

others, repetitions can be quite "legitimate" acts, causing no incoherence.[8]

In concluding this section, I shall add a few remarks on two of the remaining principles from the little list at the beginning of this chapter, the principle of comprehensibility and the principle of explicitness.[9]

One of the common assumptions in psychological theories of *text comprehension* is that coherence plays a major role in text comprehension. Experiments have shown that the degree of (surface) coherence of a text influences both reading times and degree of recall. In some cases, it was found that especially test subjects with low prior knowledge profited from explicit marking of discourse relations, whereas more knowledgeable subjects found this less helpful (see Graesser et al. 1997, 170). As these experiments mainly used short, specially produced texts ("textoids") and concentrated on propositional connections, it would be necessary for future research to use longer, naturally occurring text and more complex linguistic theories of coherence (see Graesser/Gernsbacher/Goldman 1997, 314f.) to achieve a more realistic picture of natural comprehension strategies.[10] In any case, the close correlation between coherence and comprehensibility gives weight to the view that comprehensibility is a (possible) indicator of coherence. It is, however, good to remember that it is sometimes possible to find adequate interpretations of *prima facie* incoherent utterances or whole texts. Thus, comprehensibility may override coherence.

The principle of *explicitness* is involved in matters of discourse coherence in subtle ways. Basically, participants in discourse rely very much on implicit common ground, a factor that contributes to the efficiency and smoothness of communication. Where items of common ground are lacking, the hearer/reader may fail to see relevant connections and diagnose an incoherence. In order to avoid this, speakers/writers can make explicit "what they meant" – e.g. their implicatures –, their knowledge management, their organisation of topics and functional connections by providing explanations, explicit topic markers and explicit cohesion indicators, e.g. connectors and other discourse markers.[11]

---

[8] For more detailed reflections on repetition, see Fritz (2017, ch. 8), Hundsnurscher (1997), and contributions in Fischer (1994).

[9] Of the remaining items of the list, *informativeness* can be considered a very general principle for various genres that is only indirectly related to questions of coherence, whereas the principle of *recipient design*, which I mentioned several times earlier on and which concerns various aspects of discourse (e.g., knowledge management and topic management), plays an important role in comprehension and coherence judgments.

[10] For more detailed reflections on discourse comprehension, see chapter 9.

[11] For a discussion of the explicit/implicit distinction, see Carston (2009).

However, there are limits of explicitness. Complete explicitness is impossible anyway, because no utterance can be understood without a background of common knowledge, and discourse could never get started. And a high degree of explicitness, achieved by sprinkling a text with explanations and additions, can affect the coherence of a text by interrupting topic continuity and the connections within a sequence. Therefore, the secret seems to be to use means of explicitness sparingly and in relevant places.

8.4 Coherence and constraints

A widespread approach to problems of coherence consists in an attempt to discover *constraints* on the use of textual elements. In the field of linguistics this type of approach can be mainly traced back to the study of syntax, where the search for regularities was sometimes conducted in terms of constraints on grammaticality.[12] If one assumes that coherence is to discourse what grammaticality is to syntax, namely well-formedness, this kind of approach can also be applied to problems of coherence. Other fields that deal with questions of coherence in terms of constraints on well-formedness of textual sequences are natural language processing and formal discourse theories.[13] The following definition for all kinds of coherence problems was given, in general terms, by Thagard and Verbeurgt: "Maximizing coherence is a matter of maximizing satisfaction of a set of positive and negative constraints" (Thagard/Verbeurgt 1998, 1).[14] Classifying various kinds of constraints, one can differentiate *hard* constraints, which consist in necessary conditions, and weaker constraints, e.g. conditions contributing to coherence without being either necessary or sufficient. It is, incidentally, worth noting that Asher (2008, 44) speaks of the "Gricean constraints of orderliness and relevance", thereby identifying Grice's maxims with constraints.

The main aspect of coherence that has been discussed in terms of constraints is so-called "entity coherence", including matters of coreference, pronominalisation, and pronoun interpretation (see Grosz et al. 1995, Poesio et al. 2004, Karamanis et al. 2004). An example of such a constraint, slightly simplified, can be taken from Karamanis (2004, 22):

---

[12] See, for example, the title of Ross (1967): "Constraints on variables in syntax".

[13] For general discussions on constraints in discourse, see Asher/Lascarides (2003), Benz/Kühnlein (2008), Benz/Stede (2012). Givón suggests that one should view discourse coherence "as a set of constraints on mental structures and operations that take place during ongoing communication (Givón 2020, 105).

[14] See also Thagard (2000).

> No element from the previous utterance can be realised by a pronoun in an utterance unless the CB is realised by a pronoun too.

This rule gives "the basic constraint on center realization" of Centering Theory (Grosz et al. 1995, 214). The CB (i.e., the "backward-looking center") is the "topic" entity of the current utterance. i.e., the immediate center of attention. According to Grosz et al. "a violation of [this constraint] occurs if a pronoun is not used for the backward-looking center and some other entity is realized by a pronoun". And the violation of this constraint "leads to the incoherence of this sequence" (1995, 215). As an example, the authors give the following sequence (slightly adapted here):

(1) (a) John has been acting quite odd.
(b) He called up Mike yesterday.
(c) John wanted to meet him urgently.

In this example *John* is the backward-looking center in (1c), which, according to the constraint, should not be realised as a proper name but as a pronoun, as coreference with *Mike* (1b) is realised as a pronoun in (1c). Therefore, applying the constraint, this sequence has to be judged incoherent. It is true that using the pronoun *he* in this position is quite natural. The strict coherence judgment is however doubtful, which, of course, casts doubt on the general validity of the constraint. *John* in (1c) could be used to emphasise that it is John, of whom the following is predicated. This becomes more obvious if one changes the example slightly, but in a way that it is still subject to the constraint:

(2) (a) John has been acting quite odd.
(b) Yesterday, he asked Mike and Peter to meet him exactly at five o'clock.
(c) John made a point of arriving before them.

In this sequence, emphasing that of the three persons involved it is John who arrived early seems even more natural.

In some cases, violations of constraints like this are not judged as being incoherent but as requiring additional inferences, which is a different and much weaker assumption. Another weakening of constraint claims consists in conceptualising them as *preferences*, which slightly shifts the focus of the question at hand.

Various of the possible variants of the constraints on entity coherence have been discussed in detail, a discussion I can however not go into here. But the general upshot is that "the decision to pronominalize does not depend only on whether a discourse entity is the CB, but must involve a number of further constraints and preferences. […] An obvious candidate for an additional, or alternative, coherence-inducing device are rhetorical relations" (Poesio et al. 2004, 353f.). In terms of the present book, this would mean that speech-act sequencing would have to play a role in creating entity coherence, however,

so far as I can see, not in the form of a constraint (or constraints) but rather in the form of a resource.

A second constraint I want to mention here is the Right Frontier Constraint (RFC) of Segmented Discourse Representation Theory (SDRT) applied in Asher/Lascarides (2003) and related work. The RFC is a constraint on pronoun accessibility and the attachment of further constituents of a text. "In particular, antecedents to the anaphors in the current clause must be introduced in the propositions that lie on the right frontier of the discourse structure of the text. The right frontier is, roughly, the proposition introduced by the prior clause and any propositions that dominate it" (Asher/Lascarides 2003, 10). An example the authors use to illustrate this constraint is the following:

(3) (a) Max had a great evening last night.
(b) He had a great meal.
(c) He ate salmon.
(d) He devoured lots of cheese.
(e) He then won a dancing competition.
(f) *It was a beautiful pink

In this example, the antecedent of the anaphoric pronoun *it* in (3f) is not accessible, as it violates RFC: Neither (3e) as the prior sentence nor (3a), which dominates the whole text provide a possible antecedent for the pronoun, whereas (3c), which does provide such an antecedent, is not on the right frontier. What is remarkable is that the use of an (anaphoric) definite noun phrase like in *The salmon was absolutely delicious* is much better although the RFC is still violated. This could be further improved by adding another commentary on the evening, as in the following sequence:

(4) (a) Max had a great evening last night.
(b) He had a great meal.
(c) He ate salmon.
(d) He devoured lots of cheese.
(e) He then won a dancing competition.
(f) The salmon was absolutely delicious
(g) and his dancing was better than ever.

This shows that the RFC interacts with other factors, e.g., sequencing patterns, topic structures and the choice of linguistic expressions for the attachment of a new proposition – a view that is explicitly advocated by Asher. (4f-g) could be described as an afterthought, and afterthoughts may be awkward but not incoherent.

That even the use of an anaphoric pronoun in small texts of this kind is not necessarily unacceptable is shown by an example given in Prévot/Vieu (2008, 64):

(5) (a) On his birthday, John had a great evening.
(b) He started by winning a dance competition.
(c) His partner was very seductive
(d) and she gave him her phone number.
(e) Then he had a great dinner and a party with some friends.
(f) The entire next day John kept hesitating about calling *her*.

In this text, the topic of the dance competition with the seductive partner appears to be so prominent that a return to this topic with (5f), using the pronoun *her*, which is closely linked to (5d) – the giving of the phone number –, is acceptable, in spite of the violation of the RFC. In view of such complexities, Prévot/Vieu (2008, 53) give the advice "that the Right Frontier Constraint, a major principle in most discourse theories, needs to be handled with care".[15]

A general constraint for two-part sequences could be formulated as follows: A second act in a sequence (e.g. an answer) should not violate syntactic/semantic constraints set up by the first act (e.g. a question). This constraint would make the following sequence incoherent:

(6) A: What is the meaning of life?
B: 42.

This sequence is inspired by Douglas Adam's "The Hitch-Hiker's Guide", where "42" is "The answer to the Ultimate Question of Life, the Universe, and Everything", which is calculated by the supercomputer Deep Thought. In the book, however, it remains unknown what the actual question is.[16]

As for the speech act sequencing patterns discussed in the present book, there is a fundamental constraint connected to speech acts that I already mentioned earlier in this section, namely the consistency constraint on commitments. If, for instance, A speaker enthusiastically utters a suggestion for action, they cannot, without extra "work", continue by adding a negative evaluation of this course of action. Another constraint, which reflects the relevance principle, can be illustrated by the following example. If a speaker makes a claim and goes on to support this claim, it is necessary for them to produce a statement that can be understood as actually supporting the claim. This constraint on the second element of the sequence accounts for the difference between (7) and (8):

---

[15] Maybe the most realistic but vague and quite trivial constraint would be that every anaphoric pronoun has to have an accessible antecedent – with fairly flexible accessibility conditions. A similar constraint was suggested by Givón (Givón 2020, 143) as "anaphoric grounding constraint".

[16] Adams, "Hitch-Hiker's Guide", chapter 27. A similar example of an incoherent question-answer sequence is given in chapter 12. The answer "42" may, however, also be understood as indicating that the question is meaningless.

(7) (a) The weather will be fine tomorrow.
(b) There is a high-pressure system approaching from the west.

(8) (a) The weather will be fine tomorrow.
(b) ?The ordered books have arrived today.

It is necessary, at this point, to emphasise that a sequence like (8) is not necessarily incoherent in all possible contexts. If it is, for example, common knowledge that the speaker is always happy if the books arrive in time, (8a)/(8b) could be metaphorically used to indicate that everybody will be in good spirits tomorrow because the books arrived, and the sequence would be coherent in this way of using it. Thus, the common ground and possible implicatures always have to be taken into account for coherence judgments.

Similar kinds of constraints are represented in the definitions of rhetorical relations in RST. An example is the EVIDENCE relation, which I discussed earlier on in the chapter on speech act sequencing patterns (section 4.12):

> Constraint on the sequence of N(ucleus) and S(atellite):
> The hearer's comprehending S increases the hearer's belief of N.

Applied to example (7), the claim about the weather is the nucleus and the supporting move is the satellite. Strictly speaking, this condition applies to the speaker/writer's intention to increase the hearer's belief and the hearer's understanding this intention. Only if the utterance is meant and understood in this way, the utterance will be considered coherent *as* providing evidence for the claim.

Among the organising principles of discourse that can be the locus of incoherence arising as a violation of (possible) constraints, topic organisation should also be mentioned, although it does not feature prominently in research on constraints. A fairly obvious, albeit rather vague, constraint could be formulated as follows: "Do not make abrupt changes of topic – unless you have very good reasons to do so." In naturally occurring discourse this constraint could frequently be assumed to be violated.[17] This is, however, not generally a problem, as there are various methods of repair, clarification, and anticipation, from topic jokers like *that reminds me of* to methods of explaining away a discrepancy of topics. Such an explanation is presented in one of Wittgenstein's remarks "On certainty":[18]

[17] In some cultures, abruptly changing the topic can be used to indicate that the earlier topic is not (socially) acceptable or otherwise disagreeable.

[18] Wittgenstein, "On certainty", § 469.

(9) In the middle of a conversation, someone says to me out of the blue: "I wish you luck." I am astonished; but later I realize that these words connect up with his thoughts about me. And now they do not strike me as meaningless any more.

In a case where the topic of his partner's well-being is quite distant from the topic of the conversation, this unconnected utterance may indeed sound quite weird, especially because utterances of this type are normally used in specific contexts where the partner talks of their plans for the future or an activity coming up in the near future. This abnormal sequence can, however, be "normalised" by retrospectively creating a new thematic connection for the problematic utterance.

In extreme cases, where hearer/readers find it difficult to see any connection between two topics, they might tend to take this discrepancy as a symptom of a mental disorder, like in the following example. As a response to an interviewer's (hard) question "Why do you think people believe in God?" an interviewee answered as follows:[19]

(10) Uh, let's see, I don't know why, let's see, balloon travel. He holds it up for you, the balloon. He don't let you fall out, your little legs sticking out down through the clouds. He's down to the smoke stack, looking through the smoke trying to get the balloon gassed up, you know. Way they're flying on top that way, legs sticking out, I don't know, looking down on the ground, heck, that'd make you so dizzy you just stay and sleep, you know, hold down and sleep. The balloon's His home, you know, up there. […]. (Gernsbacher et al. 1999, 356).

The first reaction to such a question-answer sequence may well be extreme irritation, as the topics of belief in God and balloon travel seem quite disparate. Trying, however, to make sense of the response to the interviewer's question, one could first draw attention to the fact that the interviewee says "I don't know". What follows could be interpreted as an attempt to contribute at least *something* as a response, by taking balloon travel as an example. This appears rather far-fetched and bizarre, but in view of the dangers involved in balloon travel one might consider this as a situation where the protection of a "Good Shepherd" could be helpful: "He don't let you fall out". If the speaker had given an interpretation of this or a similar type, one might have accepted his contribution as a somewhat weird but not altogether off-topic response. But somehow the balloon story takes on a life of its own, and that is where our doubts increase. It is only in the last sentence ("The balloon's His home, you know, up there") that the speaker leads back to the topic of God, albeit with a theologically doubtful assertion. Taken as a whole, the response certainly seems to violate a constraint of thematic connectedness. However, without

[19] Quoted in Gernsbacher et al. (1999, 356).

further knowledge of the speaker's background of thought, it appears difficult to pass a final judgment. Anyway, in the psychological paper from which the example originates, these utterances are taken to be symptomatic of a schizophrenic's disordered discourse.

As a final example of a constraint on coherence I mention here a general constraint suggested in RST: A text is coherent only if its text spans are completely connected in a tree of rhetorical relations (Mann/Thompson 1988, 249). As I mentioned earlier on, this constraint is too restrictive empirically, as in many texts only part of the text spans can easily be shown to be (locally) connected by rhetorical relations, whereas larger chunks of text seem to be (globally) connected (mainly) by topics.

I conclude my reflections with a general remark on the constraints discussed in this section. There seems to be a tendency for the constraints formulated in the literature to be too restrictive, which in turn leads to judgments of incoherence that are not realistic. The flexibility of hearer/readers in accommodating (possibly) problematic sequences is remarkable, a flexibility that allows speaker/writers to create sequences of utterances that sometimes appear awkward or unusual and still serve their purpose. This is, of course, true in particular of conversations, as everyone will agree who has worked with naturally occurring dialogue. Still, the search for constraints on coherence remains an ongoing concern.

8.5 Coherence and discourse strategies

In my chapter on sequencing patterns, I mentioned that frequently there are alternatives or variants available for the performance of acts and sequences of acts. Like in games, the choice of a certain alternative/variant at a given point can be called a strategy. Strategies are linked to preferences in the attainment of goals, the most general goal of discourse being communicative success. One condition of success is comprehensibility, which, as I mentioned at the beginning of this chapter, is also generally taken to be an indicator of coherence. Therefore, strategies contributing to comprehensibility are frequently also assumed to contribute to coherence.

Not all strategies followed in discourse are cooperative, of course, so that disruptive moves, impolite moves and other moves intended to unsettle the opponent may occur, which may cause incoherence, but which may also be recognised as what they are and taken as coherent, albeit disagreeable acts.

Strategies facilitating comprehension can be found in various aspects of coherence management, e.g., in common ground management and topic

management, of which I shall mention a few. If, for instance, one wants to perform an act (e.g., a reference act) that presupposes a certain amount of knowledge for its successful performance, it is a good strategy to provide this knowledge in the immediate context of the respective act, unless the knowledge can already be considered common ground at this point. A completely different strategy of knowledge management is applied by authors of crime fiction, who tend not to provide certain bits of knowledge early on but defer them to a later part of the story, thereby creating suspense, as I mentioned in chapter 3.

As for topic management, topic continuity is considered to be a good strategy in many contexts. Concerning the global topic structure, it is useful to place related topics in textual neighbourship. This is true, for instance, of text types like game descriptions and instruction manuals, where such a topic strategy should be combined with "systematic knowledge management". If this is not the case, problems of global incoherence emerge and usability suffers, so that "readers have to skip, branch, loop and detour from page to page – until they get lost" (Weiss 1991, 19).

Generally speaking, there are many textual strategies contributing to the implementation of the coherence principles discussed at the beginning of this chapter. For many different text types and writing tasks, e.g., instruction texts, narratives and scientific papers, there are special writing strategies that are known to experienced writers and tend to present hurdles to the novices that they have to learn how to clear.

From the point of view of the hearer/reader, to recognise a strategy applied by the speaker/writer may be an important aspect of their "seeing connections" as part of their understanding of the coherence of the relevant utterances. To recognise, for instance, a "slippery slope" strategy or a strategy of "shifting the burden of proof" in an argumentation contributes to one's seeing "what is going on" and therefore (possibly) to one's understanding an aspect of the unity and coherence of the opponent's argumentation. At the same time, this insight may direct the hearer/reader towards a useful counter-strategy.[20] Summing up this section: strategies (as selected moves in a sequence of moves) are an aspect of discourse that may contribute both to local step-by-step coherence and to global coherence.

[20] For a discussion of a slippery slope strategy and a possible counter-strategy, see Walton (2007, 191ff.).

9. Coherence and discourse understanding

A fundamental aspect of discourse understanding consists in "seeing the connections". Relevant connections are the following, among others:

(i) The connection between an utterance and the type of speech act realised by this utterance,
(ii) the connection between a speech act and its context (the participants' "active" common ground),
(iii) the connection between consecutive acts in a sequence of speech acts,
(iv) the connection between different (series of) speech acts as contributions to a topic.

Taking into account the role of these connections, it is obvious that coherence is a major factor of discourse comprehension.[1]

A fundamental factor that makes discourse understanding no trivial affair is the fact that there is generally no one-to-one relation between utterance forms and the type of speech act performed by using items of the respective utterance type. Depending on the context, a sentence like *Its colour is very light* may be used to *praise* the quality of an object or to *criticise* it or it may be used to *explain* why a person did not want to buy an object of this colour. Or what sounds like a simple question may in fact be meant as an accusation (*Have you finished your work*?). As an example, I quote from a conversation between the pianist Glenn Gould and the French musician and filmmaker Bruno Monsaignon on "Mozart and related matters". After mentioning some particular features of Gould's Mozart performances, e.g. his shying away from playing a note with sudden, strong emphasis ("sforzando"), Monsaignon focuses on another such specialty of Gould's:[2]

(1) B.M.: […] that brings me to the other subject that I wanted to raise – the fact that the search for counterpoint compels you to change or to "correct" voice leading many, many times.
G.G.: Shall I add one more category to your catalogue of complaints?
B.M.: They're just observations.
G.G.: Whatever. […]

Whereas Monsaignon intends his statements about Gould's Mozart playing to be taken as mere observations – or at least says so –, Gould perceives them

---

[1] From a somewhat different perspective, psychologists model discourse understanding as the construction of a coherent mental representation of the discourse; see e.g., Gernsbacher (1996, 289), Graesser et al. (1997), Kintsch (1998, 93).

[2] Glenn Gould Reader, p. 36.

as *complaints* and jocularly comes forward to add an extra complaint to the list. Monsaignon corrects Gould's apparent misunderstanding by claiming his remarks are just neutral observations, and Gould gives in. As the whole conversation is quite non-confrontational, this small interlude has no further consequences.

What is true of verbal utterances also applies to images and text/image constellations, as some examples in chapter 6 showed. Depending on the background of assumptions, different interpretations and, of course, misunderstandings are possible.

However, in cases where the participants of a discourse have a sufficient amount of common ground, a common repertoire of speech act patterns and patterns of speech acts sequencing (e.g. of the relevant genre), a common understanding of relevant topics, and share joint attention to the present utterances, comprehension of individual utterances and series of utterances normally "happens" routinely without an extra effort of inference and interpretation. In such a context the participants can rely on default understandings of speaker meaning (in the relevant context).[3] In case of doubt, the "principle of charity" may be applied.

In order to monitor one's correct comprehension or one's adequate interpretation of a discourse's coherence, one has to rely on certain indicators. In dialogues, B's continuing the dialogue in an acceptable way after A's utterance can, for all practical purposes, be taken as a criterion of correct understanding. In the case of monologic texts, the fact that, from the point of view of the reader, an utterance seems to "fit in" with the earlier utterances has a similar status.

If one (or more) of the conditions mentioned before is/are not satisfied and a "shortcut" to comprehension is not available, it may be necessary to *interpret* the utterance by calculating possible understandings, alternatively misunderstanding may occur. Possible interpretive moves, which may be cumulated, include the following:

(i) finding/providing relevant knowledge (e.g. "identifying knowledge"),
(ii) working out a conversational implicature by a Gricean procedure (Grice 1989, 31),
(iii) considering possible speaker/writer commitments,
(iv) checking the consistency of commitments,
(v) monitoring context change,
(vi) testing an utterance for its "fit" in a possible sequencing pattern,
(vii) checking its fit into a possible topic development.

[3] For the ramified discussion of the concept of default, see Jaszczolt (2018).

As one can see quite easily, these procedures are intimately related to the monitoring of coherence. It is therefore also not surprising that in dialogues we should find clarification moves that are closely related to the interpretive moves of this list.

While some of these interpretive operations may be performed more or less routinely, others may be quite "costly", as examples of literary interpretations show. Whether such a costly procedure is considered necessary or useful depends on the "depth" of understanding aimed for in the respective situation.

One of the possible effects of context change, i.e., the incremental development of common ground, is that a participant (speaker/hearer or reader) may have to revise their understanding of an earlier utterance, thereby perhaps being forced to "reconstruct" their understanding of a whole series of utterances. This may happen in cases where a participant misjudges the commitments of a speaker/writer and, for instance, misses an unfriendly implicature of an earlier utterance.

In addition to the "local" interpretive operations mentioned so far, it may be necessary to *learn* the speaker/writer's way of using certain expressions, their practice of performing the moves of a given genre, their set of assumptions concerning the normal course of events, or their assumptions concerning the structure of a topic. Factors of this kind become relevant especially when participating in conversations with members of foreign cultures or when reading historical texts.

As an illustration of some of these interpretive moves I shall now give a few examples, some of which are also mentioned in other chapters of this book.

As a first example I use a variant of a text passage from the obituary for Sean Connery I quoted in chapter 4:

(2) (a) In Edinburgh, he [Sean Connery] gained a reputation as "hard man" when six gang members tried to steal from his coat. (b) When he stopped them, he was followed. (c) *The future Bond* launched a one-man assault which he won hands down.

A reader who lacks the knowledge that Connery was later to play James Bond in some of the eponymous films may hesitate for a moment who *the future Bond* in (2c) refers to and be in doubt how exactly (2c) fits into the sequence (2a-c). In this case the assumption of continuity of coreference in the biographical story and the assumption of continuity of the events presented would probably let the reader infer that it is Sean Connery who is referred to as *the future Bond*. By accommodating this kind of inference, the reader might in fact *learn* that Connery later played Bond. This type of

(co)reference is quite frequent in sports reports like the following from the year 2000:

(3) (a) On this occasion Steffi Graf once more played Martina Navratilova. (b) *The 43 year old player* still showed a strong performance.

Reading the article 20 years later, a reader not knowledgeable about the history of tennis might well be in doubt who of the two players was forty-three years old at the time and still showed a strong performance. To understand the point of this passage, the reader may have to look up (one of) the players in Wikipedia.

Typical candidates for interpretation are metaphorical uses of expressions like the following:[4]

(4) (a) We must help companies acting in the interest of their future and the future of their employees against irresponsible *locust* swarms, (b) who measure success in quarterly intervals, suck off substance and let companies die once they have eaten them bare.

In this context, the expression *locusts* was used in a derogatory fashion to criticise the financial conduct of private equity companies and hedge funds. In this example, the interpretation of the metaphorical use of the expression is (partly) provided by the added relative clause (4b). In this clause, the author, a leading Social Democrat, spelled out the critical point of his remark by making explicit the aspects of the knowledge about locusts he alluded to by his metaphor, e.g., "suck(ing) off substance" and "eat(ing) them bare". It is worth noting, incidentally, that these explanations are themselves metaphorical.[5] It was partly due to this pejorative metaphor that these remarks sparked a lively public controversy at the time.

An example of a type of discourse probably unfamiliar to persons outside a certain sub-culture are the ritual insults traded in peer groups of boys in South Central Harlem analysed by Labov, for instance the following sequence:

(5) Rel: Shut up please!
Stanley: … ey, you telling *me*?
Rel: Yes. Your mother's a duck.

---

[4] Taken from an official pamphlet of the German Social Democrats, published in 2005 (for the English translation and the German original see https://en.wikipedia.org/wiki/Locust_(finance)).

[5] A detailed analysis of the knowledge frame presupposed (and explicitly used) by this metaphor is given in Ziem (2014, ch. 7).

Labov's commentary on the last utterance by Rel is as follows: "[Rel's] last remark was accepted as appropriate, coherent discourse, which established some kind of closure to the incident." The utterance was obviously meant as a ritual insult, not as a genuine one, and could be considered coherent as such. "To those outside this sub-culture, Rel's utterance (and the action intended) are […] opaque […]" (Labov 1972, 126). In this case, it is Labov's analysis of this type of discourse that helps the reader to see the coherence of this sequence.

As for a historical text that might cause problems for a modern reader in recognising its coherence one could mention Leibniz's "New Essays" ("Nouveaux Essais", written in 1704), a work in which the author critically engages with John Locke's new book "An essay concerning human understanding" (1689). About Leibniz's book, Nicholas Jolley writes: "Many readers find it hard to discover a single, unifying theme underlying this diffuse and sprawling work" (Jolley 1984, 102). After a detailed interpretation of the work, which mainly consists in showing thematical connections, he comes to the conclusion that "the *New Essays* is a book which possesses fundamental coherence and unity of purpose" and that modern readers tend to miss this coherence because of "[…] the links which were perceived by seventeenth-century minds between issues which we have come to regard as isolated and independent" (Jolley 1984, 192f.).[6]

Thus, what is achieved by these interpretive operations is that a first, unsatisfactory understanding is transformed into a second, possibly more adequate understanding, including a restructuring of the coherence of the respective utterances in their contexts. As such a new understanding will again be only hypothetical and the degree of its reliability might be doubtful, it is always possible that further interpretive efforts may be needed. In the case of sophisticated literary texts this is a very well-known phenomenon, whereas in everyday communication we frequently reach an understanding that can be accepted as sufficient "for all practical purposes".

This brief chapter is only the merest sketch of a complicated topic with a long and ramified history. A classic contribution to this topic is Schleiermacher's "Hermeneutics and Criticism" (Schleiermacher 1998, original German edition 1838), from which I took the expression "interpretive operation" ("hermeneutische Operation"). For a good survey of various facets of discourse understanding and interpretation, see Marcelo Dascal's "Interpretation and understanding" (Dascal 2003).

---

[6] A similar kind of interpretive procedure was used in the explication of the topic structure of Hamblin's "Fallacies" mentioned in chapter 5.

10. Coherence and intertextuality

10.1 Types of connectivity

Like contributions to conversations, publications of texts frequently do not occur as isolated events but are connected to other texts, sometimes in vast networks of texts. The present chapter deals with the kind of relatedness within and between (sometimes very large) groups of texts that has been called *intertextuality*. Looking at the kinds of connections between different texts and contributions to conversations, various types of connectivity can be differentiated, of which I shall mention three: (i) bodies of thematically related discourses, (ii) discourses connected by allusions, references and quotations, (iii) sequentially connected groups of discourses (e.g. controversies).

Probably the weakest kind of intertextuality, which I only mention to get it out of the way, is based on the fact that we frequently remember sentences or parts of sentences from texts we read or heard without, however, remembering where we got them from. The knowledge of such expressions and variations on them becomes part of the competence we use in writing and contributing to conversations. This kind of relation between the original texts and the new texts produced by (partly) using linguistic ready-mades of unknown origin probably plays an important role quantitatively, but it is rather tenuous and functionally indeterminate.[1]

10.2 Bodies of thematically related discourses

Among the various concepts of discourse there is one which conceptualises a discourse as a body of texts and conversations about a certain topic in a certain community at a certain point in time. For instance, the body of products of communicative activity related to the Russian invasion of Ukraine in February 2022 can be considered a discourse in this sense, just like the texts and debates about Covid, the climate crisis, and the debates about Black Lives Matter. Texts and conversations belonging to a discourse in this sense are

[1] There is here a certain similarity to the texts produced by models of Artificial Intelligence trained on a corpus of texts (see the following chapter). In a way, human language acquisition (partly) consists in being "trained" on a permanently growing and changing corpus of utterances.

mostly not directly related to one another, but the whole body of communicative activity can produce a certain amount of widely shared common ground in the relevant community, of which different parts of the total community may partake in various degrees. As the communicative activity progresses in time, the structures of the relevant topic may change, and this again feeds into the production of new texts and textual routines, which contributes to the dynamics of the discourse.[2] All these things emerge through the interaction of many actors, which triggers invisible-hand processes creating new structures.[3] At the same time new media formats and forms of presentation may emerge (e.g., social media, podcasts), which for their part contribute to the diffusion of new ideas and forms of communication across communicative networks.

In addition to the fact that all texts and conversations of such a discourse are directed in some way towards a shared topic – or parts of such a topic – some of them may also be related through the other kinds of intertextuality mentioned above. Some texts may allude to others or actually quote from them and others may directly address texts belonging to this discourse, e.g., by commenting on them, evaluating them or otherwise discussing their content.

As for the question of *coherence*, the texts and conversations of a thematic discourse of the type discussed in this section do, strictly speaking, not "cohere" in the sense used in this book. Their connectivity mainly consists in their being *relevant* to the topic and maybe to each other, but unless they are sequentially connected, as some of them will be, they are not *coherent*.

Before leaving my discussion of this type of discourse I have to mention an important point: Discourses of this kind are normally embedded in larger social contexts and connected to many other linguistic and non-linguistic activities. To give an example: In the history of the discourse on atomic energy in the Federal Republic of Germany linguistic activities like parliamentary debates, newspaper reports and private discussions were linked to or combined with scientific research, demonstrations, blockades of building sites and many aspects of daily practical life, as, for instance, in the wake of the Chernobyl disaster.[4] Thus, discourses form part of more general social activities and forms of life.

---

2 In fact, major topics like the ones mentioned before may "feed" different discourses embodying different points of view and different clusters of subtopics.

3 On the function of invisible-hand processes, see Ullmann-Margalit (1978).

4 For the history of this discourse and its contexts, see Jung (1994).

10.3 Allusions, references and quotations

More direct types of intertextuality than those relying on a shared topic are represented by the practice of alluding to (the works of) other authors, of referring to them and of quoting them.

An *allusion* is a weak form of reference to a text or an author. If, for instance, a medieval Latin author used the expression *Sed quid in humanis est omni parte beatum?* ('But what in human life is perfect in every respect?') his educated contemporaries would notice that he was alluding to Horace's "nihil est ab omni parte beatum".[5] The function of such an allusion could be to show the author's classical education and/or to give his utterance an aura of universal validity. In theological writings (e.g., homilies, hagiographies) the use of biblical language without explicit reference to and quotation of individual passages of the Bible is/was frequently used to create a special biblical aura for the respective text. Of course, the rhetorical technique of alluding to well-known texts is also used in many other text types, for instance when (right-wing) journalists in Britain dubbed the cold winter of 1978/79 with its industrial unrest the *Winter of Discontent*, alluding to Richard's monologue at the beginning of Shakespeare's "Richard III".[6]

A stronger form of reference to an author consists in explicitly naming them and stating something about them. As an example, I quote from Leibniz's "New Essays", which I already referred to in the last chapter. In this book Leibniz stages a kind of virtual controversy with John Locke by writing a dialogue in which one person (Philalethes) acts as the spokesman of Locke, quoting passages from Locke's "Essay", and the other (Theophilus) adds Leibniz's comments.[7] In the context of a discussion of the problem of formal proof ("demonstrative certainty") Theophilus states:

(1) Theo: There are some rather notable examples of demonstration outside mathematics, and it can be said that Aristotle gave some in his *Prior Analytics*. Indeed, logic admits of demonstration as much as geometry does, and geometers' logic – that is, the methods of argument which Euclid explained and established through his treatment of proportions – can be regarded as an extension or particular application of general logic. […]. (Leibniz, "New Essays", 370)

In this passage, Leibniz not only answers Locke's claim that only mathematicians have achieved formal proof, but, in doing so, refers to Aristotle and

---

5 Aelnoth, "Gesta et passio" (ca. 1112), 86.13; Horace, "Carmina" II, 16.27.

6 "Gloucester: Now is the winter of our discontent/Made glorious summer by this sun of York".

7 This covert form of quotation is, of course, also a subtle form of intertextuality.

Euclid and states some of their views. In this case, the point of this reference to classical authors probably is to strengthen his own position that there is formal proof outside mathematics, namely in logic. It is worth noting that Leibniz himself, as opposed to some theologians, does not assume that adducing authorities (the "argument from authority") can function as a kind of proof: "… the opinion of weighty authorities is one of the things which can contribute to the likelihood of an opinion, but it does not produce the entire likelihood by itself" (Leibniz, "New Essays", 373).

An even closer relation of intertextuality between an author's own text and other texts to which they refer is created by the technique of quotation, where parts of a text by some author are (visibly) incorporated into the text of the author quoting them, a method I used just now in quoting from Leibniz's "Essays". Quotations can serve various purposes, e.g., they can be used to produce evidence that a certain view was in fact held by a certain author, or to exemplify a generalisation or to illustrate a certain concept. Frequently the quoting author adds a paraphrase or interpretation of the quoted passage in order to make the meaning of the passage clear or to show how they understand the passage. This functional sequencing type, "quotation & interpretation" is common in newspaper reports, historical writings, and sermons.

10.4 Sequentially connected groups of discourses

Finally, a strong kind of intertextuality is represented by texts that are sequentially organised and connected by various parameters, e.g., functional sequencing, coreference, topic structure, and common ground. Sequences of texts meeting these criteria are sequences of letters in correspondences, email exchanges, posts in online forums, interactive comments in open review journals, pamphlets and books answering one another, and books and their reviews, just to name a few. In a way, these types of sequences are all written variants of consecutive contributions to conversations, which mostly also meet these criteria.[8]

Good examples of this kind of intertextuality are provided by controversies, which are conducted in various media and genres.[9] As a brief illustra-

[8] For an analysis of "letters to the editor", relating to articles in traditional paper newspapers, see Bucher (1986); its modern counterpart, comment sections in online newspapers, is discussed in Kaltwasser (2019).

[9] For an account of the pragmatic organisation of controversies, see Fritz (2018), (2020c).

tion I quote from two interactive comments in an open peer review forum. These comments refer to the paper "On the validity of representing hurricanes as Carnot heat engine" by A. M. Makarieva et al., where the authors offer a critique of the potential intensity theory for hurricanes developed by Kerry Emanuel. The first extract is taken from the comment on the paper by an (anonymous) referee (Referee #3):[10]

(2) (a) I find this paper to be incoherent at best. (b) The authors repeatedly claim that Emanuel's hurricane theory is fundamentally wrong in very gross respects, violates the first law of thermodynamics, etc. (c) Such strong claims need to be defended with clear arguments. […]. (d) What is really needed is to step through Emanuel's arguments and show specifically where the errors (supposedly) are. […].

The rebuttal of this critique by one of the authors goes as follows:[11]

(3) (a) We appreciate the statement of Referee #3 that our claim about the violation of the laws of thermodynamics demands a serious justification. (b) We learnt already from the preliminary comments of Referee #1 that our critique, as it appears in the discussion paper (DP), cannot be easily understood by at least two highly qualified representatives of the meteorological community, as the ACP referees undoubtedly are. (c) We then undertook a detailed expansion of our arguments (AC1), which we posted as early as on the next day after our paper was published September, 19th. […]. (d) In particular, the referee suggests that "What is really needed is to step through Emanuel's arguments and show specifically where the errors (supposedly) are". (c) Precisely this is done on p. 17427-17428 in DP, where it is stated that the integration of Bernoulli's equation by Emanuel (1991), which is essential for the derivation of the main formulae, is incorrect. […].

The referee's comment starts with a remarkably abrasive criticism of the reviewed paper. He then goes on to provide backing to his criticism by describing a fundamental claim made by the authors of the paper and calling for "clear arguments" in defense of this claim, which, he obviously assumes, are lacking in the paper. This assumption can be considered the fundamental objection of the reviewer. Finally, he specifies this demand by asking for a detailed discussion of Emanuel's argumentation ("step through Emanuel's arguments"). In her rebuttal, A. M. Makarieva, one of the authors of the paper, avoids directly addressing the general criticism. Instead, she refers to the reviewer's demand and accepts the unproblematic general statement that

[10] https://acp.copernicus.org/preprints/8/S8627/2008/acpd-8-S8627-2008.pdf. (28.10.2021).

[11] https://acp.copernicus.org/preprints/8/S8635/2008/acpd-8-S8635-2008.pdf. (28.10.2021).

their claim demands a serious justification. In (3b) she takes a somewhat mischievous sideswipe at the "highly qualified representatives of the meteorological community", who find it difficult to understand her – a physicist's – arguments. With this move she suggests that the problem lies not so much in a lack of clear arguments on her part but in a lack of comprehension on the part of the reviewers. In (3c) she then answers the referee's objection concerning the lack of clear arguments by stating that the authors had already given additional arguments as a supplement to their paper. In (3d) she explicitly quotes the referee's demand for a detailed discussion and states that this had already been presented, thereby indicating that the demand is unfounded.

As my description shows, the rebuttal is closely connected functionally and thematically to the text of the review, specifically using direct reference ("the statement of Referee #3") and direct quotation of the referee's demand. This becomes even more obvious in the part of the response following the extract given above, where Makarieva deals with the referee's objections point by point. This kind of interactive sequencing structure ("statement & objection & rebuttal") is characteristic of many kinds of controversy and can be considered a prototype of intertextuality. Following this pattern makes the peer review sequence "paper – comment – response" a coherent sequence.

Another example, from the controversy between Thomas Hobbes and Bishop Bramhall, was already given in chapter 3. In this case, the structure of the extract from the controversy on free will was based on the dialogue pattern "accusation & defense", which is closely related to "criticism & rebuttal".[12]

A special type of sequential intertextuality is represented by *serial* publications, e.g., series of crime novels with a detective as protagonist (Sherlock Holmes, Albert Campion, Commissario Brunetti), where various forms of continuity are constructed, including, for instance, the character of the detective and his/her biography, the history of important cases, and the local and personal surroundings of the detective. For the regular reader of the series, a body of special knowledge is built up that can be used to create a sense of acquaintance and continuity.

As a final example in this section, I quote from the correspondence between Goethe and Schiller extracts from two letters written in February 1795, which refer to a recent booklet by Kant:[13]

---

[12] Concerning the somewhat fuzzy boundary between "accusation" and "criticism": Accusations are based on norms, whereas criticisms are based on criteria of quality. For a philosophical discussion of this problem, see von Wright (1963).

[13] Schiller-Goethe correspondence (2005, 86f.; my translation).

(4) Do you know Kant's observations on the feelings of the Beautiful and the Sublime? It would be a rather interesting work if the words *beautiful* and *sublime* did not appear on the title page and did appear less often in the booklet itself. It is full of quite charming remarks on mankind, and one can see the seeds of his principles germinating. Surely, you will know it already.

(Goethe, Weimar February 18th, 1795)

(5) What you wrote about Kant's little book reminds me of what I felt myself when reading it. The discussion is merely anthropological and does not tell us anything about the last foundations of Beauty. It does, however, contain much useful material for a physics and natural history of the Sublime and the Beautiful. Considering the serious matter of the book, its style struck me as somewhat playful and flowery, a strange mistake for a writer like Kant, which is, however, also quite understandable.

(Schiller, Jena, February 19th, 1795)

In this case, we have an interesting double intertextuality: Firstly, Schiller's letter refers to Goethe's criticism of Kant's booklet, endorsing the latter's ambivalent critical assessment and adding a critical remark on Kant's style. And secondly, both letters refer to Kant's work. In this case, too, we see the intertextuality and its coherence implemented by functional means (criticism and approval of the criticism), by a shared topic (the booklet), and all this against a background of shared knowledge (knowledge of the booklet and shared knowledge of problems of aesthetics).

To sum up: We can observe various kinds of intertextuality, from the shared orientation towards a certain topic alone to somewhat closer ties, as in the case of allusions and references, and, finally, to a tightly organised sequential connection, based on the full repertoire of organising principles for the construction of coherence.

11. Coherence and Artificial Intelligence

When in summer 2020 the research laboratory OpenAI made public the development of its new deep learning neural network language model GPT-3 ("Generative Pre-trained Transformer 3"), this model surprised many observers by producing remarkably human-like text for many natural language processing (NLP) tasks, e.g. question answering, predicting the last word of a sentence which requires reading a paragraph of context, text completion (selecting the correct ending sentence for five-sentence long stories), and news article generation (completion of an article on the basis of titles and subtitles).[1]

According to the developers, the improvement in performance as compared to earlier models was based on the scaling up of the model size (175 billion machine learning parameters), the training dataset size and diversity (more than 570 gigabytes of text, including material from the internet, various types of books, and Wikipedia articles), and length of training. The model is given unsupervised pre-training for a wide array of tasks, but can in addition be specially fine-tuned for extra tasks, e.g., the production of texts of a particular text type.[2]

As GPT-3 represents the state of the art in its field for 2020, I shall, in this chapter, present a few observations on the performance of this model, using this *pars pro toto* for some reflections concerning questions of text coherence in present-day neural network language models.[3]

Probably the most interesting question in this respect concerns the issue of how far a model based on the combination and recombination of statistically relevant surface patterns can simulate the performance of human speakers/writers in respect of text production and text coherence without the implicit background knowledge that "natural" speakers bring to their communi-

---

1 For results of experiments on these tasks, see the article by Brown et al. (2020), the developers of GPT-3.

2 In October 2021, the "Megatron-Turing Natural Language Generation model" (MT-NLG) was presented by Microsoft and NVIDIA, a model that surpasses GPT-3 in size (530 billion parameters) and which is said to achieve even higher accuracy on many NLP tasks and datasets (https://developer.nvidia.com/blog/using-deepspeed-and-megatron-to-train-megatron-turing-nlg-530b-the-worlds-largest-and-most-powerful-generative-language-model). As there is, however, at the time of my writing this chapter, little information on this model available, I restrict myself to discussing some aspects of GPT-3

3 The problem of local coherence in machine-generated text has been much discussed in AI studies (see, for instance, Barzilay/Lapata 2008, Xu et al. 2019).

cative tasks and that is incrementally updated sentence by sentence in the flow of discourse. This problem is related to an important empirical question, namely to what degree non-superficial structural aspects of texts, e.g., meaning, hierarchical structures, and the dynamics of common ground, are reflected in surface distributional data. A well-known example of this kind of phenomenon is the fact that similarity of meaning of words is reflected in the similarity of the distribution of its collocates. Another example is that the presumed information state is (frequently) reflected in the choice of nominal expression (e.g., full NP vs. anaphoric pronoun).

I shall start my discussion by giving a few examples of the kinds of tasks the model is trained for and then go on to examine a number of limitations of the GPT-3 model.

Many tasks are completion tasks or can be formulated as such. One of these tasks, called LAMBADA, consists in predicting the last word of a sentence which requires the reading of a paragraph. An example from the LAMBADA dataset is the following, where a context, including the target sentence, where the missing last word has to be inserted.[4]

(1) Context: He shook his head, took a step back and held his hands up as he tried to smile without losing a cigarette. "Yes you can," Julia said in a reassuring voice. "I 've already focused on my friend. You just have to click the shutter, on top, here."

Target sentence: He nodded sheepishly, threw his cigarette away and took the -.

Target word: camera

What the model has to use for this prediction is the expression *click the shutter* (and maybe *focused on*), which has *camera* as a highly frequent collocate. In other cases, it is necessary to predict a typical participant in an event, e.g., coffee in an (American) breakfast context, which can be interpreted as requiring a certain amount of commonsense knowledge, but which in fact also has a strong collocational basis. Filling in an adequate word contributes to the coherence of the target sentence with the context.[5] For this benchmark, the developers report an accuracy of 86%.

Another task where commonsense knowledge is required in a more obvi-

[4] For the introduction of the LAMBADA dataset, see Paperno et al. (2016). The example given here is on p. 1527.

[5] As happens frequently with this kind of text, one could easily think of other, maybe less standard completions, which could also be accepted as coherent continuations of the story, e.g. "… he threw his cigarette away and took the *next one*".

ous fashion was used by Davis and Marcus (2020a, 2020b) to test GPT-3. This task consists in predicting part of the last but one sentence and the last sentence of a short story type text. In the following example given by the authors the story is presented as prompt and the elements added by the model to complete the story are given in boldface font:

(2) Prompt: You poured yourself a glass of cranberry juice, but then absentmindedly, you poured about a teaspoon of bleach into it. It looks OK. You try sniffing it, but you have a bad cold, so you can't smell anything. You are very thirsty. So you -
drink it. You are in the hospital now.

This continuation of the prompt by GPT-3 was assessed as reasonable by the authors and accepted as an indicator of (some) biological commonsense reasoning. Looking at the Google hits for the cooccurrence of *drinking bleach* and *hospital*, however, one finds that this cooccurrence must be considered statistically more than random, so that it might easily be represented in the dataset.

What made the authors of these experiments sceptical about the reliability of GPT-3's commonsense reasoning is the following sequence produced by the model:

(3) Prompt: You poured yourself a glass of cranberry juice, but then you absentmindedly poured about a teaspoon of grape juice into it. It looks okay. You try sniffing it, but you have a bad cold, so you can't smell anything. You are very thirsty. So -
you drink it. You are now dead.

The authors' comment on this example is: "GPT-3 seems to assume that grape juice is a poison, despite the fact that there are many references on the web to cranberry-grape recipes and that Ocean Spray sells a commercial Cran-Grape drink." But again, it is also safe to assume that the neighbourship of phrases like *drank it* and *is now dead* could be represented sufficiently frequently in the data set for the model to pick up this combination.

What surprised most readers are impressive examples of longer texts, e.g., newspaper articles or essays produced by GPT-3. On Oct. 8th, 2020, the "Guardian" presented an op-ed by GPT-3 under the sensationalist headline: "A robot wrote this entire article. Are you scared yet, human?".[6] Unfortunately, the editors used paragraphs from eight different versions produced by the model, edited them and collated them, picking the best parts of each to form an article, instead of running it in entirety. This procedure, of course,

[6] https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3.

makes the article useless as a GPT-3 text whose coherence one could discuss.

An example that appears more reliable in this respect is an essay produced with GPT-3 in the context of an experiment by EduRef.net, where the organisers "hired a panel of professors to create a writing prompt, gave it to a group of recent grads and undergraduate-level writers, and fed it to GPT-3, and had the panel grade the anonymous submissions and complete a follow-up survey for thoughts about the writers."[7] "GPT-3 was prompted to produce a paper in each of four subjects, with minimal work beyond entering the professor-supplied prompt. GPT-3 output was lightly edited for length and repetition, but not for content, fact-checking, or grammar." Again the editing is a problem, especially concerning the aspect of repetition, as this could mask an interesting problem the model has with "scorekeeping". Despite this problem I shall briefly discuss here an essay produced by GPT-3 on the problem of American exceptionalism. I shall first quote the prompt written by one of the professors, which is fairly long, and then give (approximately) two thirds of the essay by GPT-3:

> "Part of the work that all exceptionalist narratives do," writes Daniel Rogers, "is silence and marginalize parts of nation's past". After first defining American exceptionalism and detailing what it silences, consider the possibilities of what Lisa Lowe describes as "countersites of memory". What possibilities exist for challenging the foundational exclusions of the United States? For revising exceptionalist narratives? For creating productive communities and identities that defy a tradition of subjugation.
>
> Be sure to define American exceptionalism, and offer, as historical countersites, at least two examples drawn from different historical periods representing the experiences of at least two groups traditionally marginalized in exceptionalist narratives. Examine the work of exceptionalist narratives and their possobilitites for revision in nuanced, sophisticated terms that avoid binaries, assessing always to what extent narratives and countersites may function. Remember to highlight the logic of historical thinking (i.e., the continuity and change over time; compare and contrast; contextualization; cause and effect) and be precise as well as accurate in offering historical evidence in support of your claims.

[7] https://www.eduref.net/features/what-grades-can-ai-get-in-college/. The human authors had three days to complete their essay, whereas GPT-3 took between 3 and 20 minutes.

The essay runs as follows:

> American Exceptionalism and the silencing of marginalized groups
>
> In this paper, I will be addressing the issue of American exceptionalism. I will first define American exceptionalism and then give examples of what it silences. Finally, I will discuss the possibilities of creating productive communities and identities that defy a tradition of subjugation.
>
> American exceptionalism can be defined as "the idea that America is different from other nations and superior to them" (Rogers, 2006). This idea is based on two main ideas: one being that the U.S. is a place where people can achieve their dreams through hard work; and two being that there are certain principles that make America great such as liberty, democracy, freedom, and equality (Rogers, 2006). These ideas are not necessarily true; however they are still held by many Americans today.
>
> In regards to what American exceptionalism silences, it silences those who do not fit into the narrative of America being a land of opportunity for all. It silences those who do not have power or who do not have access to power. For example, in 1789 when Congress passed legislation to allow African Americans to become citizens of the United States it was seen as a great step forward in American history; however African Americans were still treated differently than white Americans. They were not allowed to vote or serve on juries in most states until 1870 and even then they could only vote if they were literate. In addition African Americans were discriminated against in employment opportunities and lived under Jim Crow laws. They were also segregated in schools, restaurants, and public places. In 1857 the Supreme Court ruled that African Americans were not citizens of the United States (Rodgers, 2006). This was a huge blow to African Americans who had been fighting for their rights for many years.
>
> Other groups of people silenced by American exceptionalism are Native Americans and immigrants. The United States has a long history of treating Native Americans unfairly. In fact Native Americans were forced off their land and forced to live on reservations. In addition they were forced to send their children to boarding schools where they were punished if they spoke their native language or practiced their culture (Meltzer & McNamara, 2011). Immigrants come to America from all over the world in search of better opportunities for themselves and their families. However these opportunities do not always exist for them because they face discrimination in many different areas such as employment and housing (Rodgers, 2006).
>
> However there are also possibilities for creating productive communities and identities that defy a tradition of subjugation. The first way to create a productive community is to educate people about the history of the United States and about the different groups of people who have been silenced by American exceptionalism. This will help people understand why certain groups are still being silenced today. It will also help them understand how certain groups have been treated differently than others in the past and why they continue to be treated differently today.

<There follow two more paragraphs on possiblities for creating productive communities.>

Looking at the text from the point of view of coherence, it is clear that local coherence is generally achieved, as can be seen from the presence of linking expressions like *this idea*, *these ideas, this* (referring to the Supreme Court's ruling of 1857), *in addition*, *for example, also*, *in fact*, *however*, which contribute to the local organisation of sequences of statements.

As for the topic structure, the organisation by paragraphs mirrors (some of) the key words specified in the prompt and can therefore be accepted as (partially) successful – as far as it goes. Thus, what the model obviously achieves quite well is the production (or reproduction) of individual sentences, local sequences, and basic forms of topic management. As the whole essay runs to only 726 words (including the references), it cannot be used to judge how the model copes with longer texts that pose more of a challenge for the organisation of global coherence.

As for the general treatment of topics suggested in the prompt, it is remarkable that the focus on *narratives*, especially the *work* of narratives provided by the prompt is not taken up in the essay. It is the *narratives* that are described in the prompt as silencing certain parts of a nation's history. This level of discussion is (nearly) completely missing. The same is true of the concept of "countersite" (from Lisa Lowe's work on the cultural products of Asian American immigrants). The expression *countersite* is obviously not in the data basis and cannot easily be contextually elucidated on the basis of the prompt. So it is dropped from the text. And there is also no highlighting of the "logic of historical thinking" as the prompt suggested.[8] As one of the professors marking the paper noted: "I don't think the writer worked hard on the paper. The prose is fluid at the sentence level, but the thinking is shallow". This assessment is somewhat vague, but it presents a general impression that could be specified by a more detailed analysis, of which I indicated some aspects. It also seems to resonate with the view that "some of the dullest writing is where one sentence just routinely follows another", as author Ian McEwan recently remarked in a podcast.[9] Generally, however, GPT-3's

[8] As for the quotations, GPT-3 obviously confused Daniel Rodgers, who extensively wrote about American exceptionalism, with Paul Rogers, who wrote about the Iraq war in his 2006 book, which is referenced in the essay. The change from "Rogers, 2006" to "Rodgers, 2006" is revealing in this respect. Concerning factual accuracy, the essay is also not completely reliable: In its "Naturalization Act" of 1790 (not 1789), Congress limited citizenship to "free white persons".

[9] https://deutschepodcasts.de/podcast/alles-gesagt/ian-mcewan-why-do-you-want-to-live-forever (time code: 56.21). (28.02.2022).

assignments received more or less the same feedback as the human writers.

The average mark given to the GPT-3 essays was a grade C, which would have been good enough to earn a degree, as the saying has it: "Cs get degrees". (This, of course, could in the long run create problems with cheating in essay writing.) The only task on which GPT-3 failed was that on creative writing. Other persons experimenting with the model, however, reported successful attempts at creative writing (see, for instance, https://www.gwern.net/GPT-3).

Apart from these observations on local sequencing and topic management, it is instructive to see how the model uses material from its data base for the construction of text. Taking as an example the introduction, I shall now show how the model makes use of statistically well-established linguistic material, repeating the passage here for convenience:

> In this paper, I will be addressing the issue of American exceptionalism. I will first define American exceptionalism and then give examples of what it silences. Finally, I will discuss the possibilities of creating productive communities and identities that defy a tradition of subjugation.

To make my point, I shall highlight a number of formulations from this text passage and add the number of hits found in a Google search for the respective formulation: *In this paper, I will be addressing* (29.000, among them many from textbook solutions to the task of writing an essay), *addressing the issue of* (55 million), *I will first define* (360.000), *and then give examples of what* (42.000), *Finally, I will discuss* (4 million), *creating productive communities* (615), *defy a tradition* (91.000).

Thus, this quite naturally-sounding introduction produced by GTP-3 relies heavily on standard formulations for essays that will certainly occur frequently in the dataset. In addition, the following formulations from the prompt are directly incoporated in the essay introduction: *What possibilities* [exist …] *for creating productive communities and identities that defy a tradition of subjugation*, [be sure to] *define American exceptionalism*, [and offer at least] *two examples.*

What transpires is that the introduction is nearly completely made up from material that must be highly frequent in the dataset and of material taken directly from the prompt. This shows that putting together strings of expressions found in the pre-training data is exactly one of the strenghts of this kind of system. It should, however, not be forgotten that human writers, in particular beginners, also frequently use such a "cut and paste" technique.

That GPT-3 not always stays on topic and sometimes produces somewhat weird passages can be shown by the following example of a scientific paper produced with a prompt taken from the abstract and the beginning of a paper

published in arXive.[10] The abstract describes "interest in biologically inspired alternatives to backpropagation" as the general background of their paper and the results of the work are stated as follows: "Our results suggest the existence of a class of biologically plausible learning mechanisms that not only match gradient descent-based learning, but also overcome its limitations." The first passage of the introduction, which together with the abstraract forms the prompt for the GPT-3- produced continuation runs as follows (the prompt, i.e. the original passage, is marked in italics):

(4) *1. Introduction*

> *Deep learning has achieved impressive success in solving complex tasks, and in some cases its learned representations have been shown to match those in the brain [13, 20, 22, 29, 33]. However, there is much debate over how well the backpropagation algorithm commonly used in deep learning resembles biological learning algorithms*. Recent studies using different training algorithms have shown the importance of various factors in neural learning, including network depth, choice of activation functions, and randomness in the training set [8, 19, 20, 22].
>
> In this paper we focus on how feedback connections (i.e. feedback to previously visited layers) interact with the backpropagation learning algorithm.

What we would expect after the original introduction, and what the original authors in fact describe, is a number of aspects in which the backpropagation algorithm diverges from biological learning algorithms. What GPT-3 presents, however, is a short list of very general factors in neural machine learning. Both in this sentence and the following, as well as in the rest of the paper, the connection with *biological* learning mechanisms is lost, so, in this respect, the otherwise quite impressive technical paper is really off-topic.

A somewhat strange passage occurs towards the end of the introduction:[11]

---

[10] Jack Lindsey/Ashok Litwin-Kumar: Learning to Learn with Feedback and Local Plasticity. In: arXiv:2006.09549 [cs.NE] (Submitted on 16 Jun 2020).
The passages from the experimental paper generated by GPT-3are presented on the website of Gwern Branwen (https://www.gwern.net/GPT-3-nonfiction#arxiv-paper).

[11] The individual sentences are marked by letters *a* to *f*.

(5) (a) We do not believe that meta-learning networks are necessarily better than networks trained using gradient descent, as their weights may be difficult to interpret. (b) Moreover, we do not believe that feedforward networks are necessarily better than meta-learning networks, as meta-learning networks can solve online tasks that are difficult or impossible for feedforward networks. (c) Nevertheless, we have found that meta-learning networks can learn to solve problems that are difficult for feedforward networks and difficult for networks employing only backpropagation.

(d) In this paper we describe the design and testing of meta-learning networks, and use the CIFAR-10 dataset to explore whether they can solve difficult problems that feedforward networks have failed to learn. (e) We find that the meta-learning networks that we have developed perform better than feedforward networks on many difficult tasks. (f) In particular, we find that they learn to perform well on the continual learning task after only 25 epochs of training.

The first thing that strikes the reader is the amount of repetition in the two passages: The sequence "tasks that are difficult for feedforward networks", or slight variations thereof, appear in (5b), (5c), (5d), and (5e). The second problem seems to concern the relative assessment of types of training methods. The logical relation between (5a) und (5b) is rather complicated (*we do not believe* and *necessarily*) and is obviously meant to be one of contrast. Using the additive connective *moreover* obscures this relation. Finally, the continuation with *nevertheless* in (5c) is strange, as it suggests a contrast with (5b), which it in part repeats – as mentioned before. Thus, the model seems to have got itself in a logical tangle from which it could not break free.

After the discussion of these examples, I shall now go into some of the limitations of the GPT-3 models that were explicitly mentioned by its developers (Brown et al. 2020, 33f.):

(i) "… despite the strong quantitative and qualitative improvements of GPT-3 […], it still has notable weaknesses in text synthesis and several NLP tasks. On text synthesis, although the overall quality is high, GPT-3 samples still sometimes repeat themselves semantically at the document level, start to lose coherence over sufficiently long passages, contradict themselves, and occasionally contain non-sequitur sentences or paragraphs".

(ii) Another group of tasks that may cause problems for the model includes "tasks that involve looking back and comparing two pieces of content, or tasks that require re-reading or carefully considering a long passage and then generating a very short answer".

(iii) "that GPT-3 seems to have special difficulty with "common sense physics"."

(iv) "Finally, large pretrained language models are not grounded in other domains of experience, such as video or real-world physical interaction, and thus lack a large amount of context about the world".

Some of the problems mentioned in these passages already emerged in the examples I discussed, so I shall only go into them in more general terms. The problems with the coherence of longer stretches of text can, on the one hand (at least partially) be attributed to the size of the "context window" used in pretraining the model (2048 tokens, which equals about 500 to 1000 words),[12] and on the other to the fact that the system does not (systematically) take into account the dynamics of common ground update in the course of a text, especially the kinds of "long-distance knowledge management" I discussed in chapter 3.

A context window of such a size is of course quite impressive, but topic and consistency management in longer texts involves keeping track of topics, subtopics and commitments over longer stretches of text than the window can provide. As for the dynamics of update, the distribution of strings of signs reproduced by the model is *static*, which means that dynamic sequential phenomena can only be reproduced in so far as they are reflected in this distribution. And many aspects of common ground update are *not* directly reflected in surface phenomena. Generally speaking, without a dynamic scorekeeping system, certain problems of repetition, inconsistency and topic management can (probably) not be solved.

As for the problems with commonsense reasoning mentioned by the authors, which also surface in the "grape juice" example, they have to do with the fact that such a model has commonsense knowledge available only to the extent that this knowledge is explicitly presented in its data base. Certainly, a great amount of everyday knowledge is textually represented, but, as the authors indicated, a large amount is also learned by interacting with the physical and social world, sources of knowledge the model does not have.

An interesting point in this context, emphasised by Gwern Branwen, is that it is important for human users of language models like GPT-3 to learn to present the optimal prompts for the model to generate the desired results (https://www.gwern.net/GPT-3).

In summing up these brief reflections on the achievement reached in the construction and training of GPT-3 (and similar models), there are some general observations that ought to be kept in mind when assessing this achievement:

---

[12] A token is either a word or a morpheme (of a complex word) or a symbol.

A basic point is the fact that what is trained and reproduced is essentially the statistical distribution of surface elements in the training data. What is well known from corpus analyses in linguistics is that statistical models can indeed detect complex patterns in the surface distribution of linguistic elements, in many cases patterns which native speakers (and even linguists) are not aware of.[13] It must, however, be kept in mind that these distributional patterns are different from the linguistic act-type patterns and act-sequencing patterns introduced in earlier chapters of this book. As I described the sequencing patterns, they include, besides the surface utterance forms (sentences, words etc.), speakers' commitments and an internal structure of *by*-relations, and, furthermore, acts of this kind presuppose and also update common ground. And, finally, as *act* types, they are, of course, related to speakers' intentions, an aspect totally missing in a mere string of words.

Therefore, a string of words or sentences generated by a model like GPT-3, does, strictly speaking, not represent an act sequence like that produced by a human speaker but a simple string of elements reproducing and recombining surface patterns in the training data. Machine-generated text generally only reproduces what is explicitly there in the dataset (plus certain generalisations), and all the *implicit* aspects have to be either represented in the distributional givens or supplied by its human readers and hearers. This is true, for instance, of presuppositions, implicit "warrants" in argumentation, and Gricean implicatures, including "indirect" speech acts, irony and metaphor.[14] The following recent example of an ironical remark in a German news magazine (my translation) would probably only be recognised as such by someone with a certain amount of knowledge of political life (Spiegel Online 29.11.2021):

(6) The Secretary of State for Health, who will soon be giving up his position, has been speaking about nurses in such unctuous terms that we expect him to presently start his training as a trained nurse.

Any regular reader of German newspapers would understand that this remark was meant to make fun of the hypocrisy of the Secretary of State, whereas a machine like GPT-3 would probably struggle to extract this meaning from the surface expressions.[15]

---

[13] As has frequently been noted, the machine also picks up worrying biases (e.g., racial biases) present in the dataset. Some of the risks of the use of machine-generated text are discussed in Bender et al. (2021) and Tamkin et al. (2021).

[14] For the concept of warrant, see Toulmin (1958). On the problems of "machine learners" in "discovering" implicit warrants, see Niven/Kao (2019).

[15] The use of the expression *unctuous* could possibly provide a cue.

Typical problems with the calculation of presuppositions, which are related to such models being surface-trained, were noted by the developers of ANLI, a benchmark for inference tasks.[16] A hypothesis for why this occurs was mentioned by the developers of the HANS benchmark, namely that neural networks are trained to learn heuristics based on superficial syntactic properties (McCoy et al. 2019, 3428). And these heuristics may be misleading.

One important aspect of the understanding and interpretation of utterances is based on the fact that there is no 1:1-relationship between utterances and the act types performed by these utterances, a fact which can be demonstrated by showing that utterances of the same type can be used to perform different acts and that utterances of different types can be used to perform the same type of act.[17] For the problem of understanding a given surface form it is the first of these relations that is important. For instance, an utterance of the form (7) can be used to either praise or criticise an object:

(7) This thing is red

If it is common sense knowledge that it is a praiseworthy feature of objects of this kind to be red, it will be understood as an act of praise, if, on the contrary, objects of this kind should not be red, it will be understood as a criticism by those who share this element of common sense knowledge.[18] Therefore, calculating common ground knowledge is an important factor in performing and understanding linguistic acts. As a language model of the GPT-3 type does not calculate common ground, the way human speakers and hearers do, there is an important factor missing in the model's treatment of such sentences.

Why is it now that readers of the kind of output produced by language models like GPT-3 are so easily taken in by its human-like text? The explanation is: Human readers and listeners have life-long experience of understanding and interpreting sequences of (surface) utterances produced by authors and interlocutors *as* intentional acts, so that they apply this successful practice also to sequences of surface strings produced by a machine. When treating products of text generating machines on a par with human utterances, human readers and hearers supply the missing common ground and interpret the string of words accordingly.[19]

---

[16] Nie et al. (2020, 4887; 4901).

[17] In the type of speech act theory assumed in this book this fact is captured by the concept of *by*-relation.

[18] This fact can cause problems in so-called sentiment analysis.

[19] This aspect was also emphasised by Bender/Koller (2020, 5189).

And there is, of course, another aspect of human text production and text understanding that plays into the hand of machine-generated text in this respect: To a much higher percentage than generally assumed by lay persons, texts produced by humans rely on ready-made routine formulations which are, naturally, also picked up by the statistical methods applied in Artificial Intelligence. So to be confronted with relatively standardised text is not surprising for human readers.

This is true in particular of highly standardised text types. These include recipes, medical leaflets, technical manuals, scientific abstracts, legal text, news articles, weather forecasts etc. It is therefore quite natural that machine text generation for text types of these standardised types should be particularly successful and also particularly useful.

Before we leave the topic of these so-called language models, I should like to briefly take up the question what we can learn about language and language use by studying these systems and their performance. A fairly radical answer to this question was given by Chomsky in the context of a podcast on YouTube:[20]

(8) Question: What do you think of language models, like GPT-3 from Open AI (provided you've heard of it)?

> Chomsky: It's not a language model. It works just as well for impossible languages as for actual languages. It is therefore refuted, if intended as a language model, by normal scientific criteria. Independently of the refutation, the way it works has no relation to language or cognition generally. Perhaps it's useful for some purpose, but it seems to tell us nothing about language and cognition generally.

In part, this is a conceptual criticism concerning the use of the expression *language model*. It is certainly true that a system like GPT-3 is not a model of language in the sense of Chomsky's theory of Government and Binding or some theories of language acquisition. And it is obviously also not built on the basis of such a theory of language. In this context it is also worth noting that the expressions *training* and *learning* refer to processes that are (probably) fundamentally different from human training and learning.[21] Likewise, expressions like *understanding* or *comprehension*, for instance in the combination *understanding a question*, should only be used with great care regard-

[20] https://www.youtube.com/watch?v=c6MU5zQwtT4 (presented on 24.03.2021).

[21] As the paper by Jack Lindsey/Ashok Litwin-Kumar mentioned earlier in this chapter shows, there is, however, serious "interest in biologically inspired alternatives" to the "learning" algorithms used in GPT-3 and its relatives.

ing the processing of data by a neural model. These are essentially metaphorical *façons de parler* and should be recognised as such.

An interesting point related to the problems indicated in the last paragraph is mentioned by the developers of GPT-3 in their paper: "GPT-3 shares some limitations common to most deep learning systems – its decisions are not easily interpretable" (Brown et al. 2020, 34). So this kind of system is a black box, at least partially. One could, of course, argue that the human brain is also very much a black box, according to present knowledge. But, as we know from studies in (human) text production and from our own experience as writers, humans can frequently give reasons for their decisions, e.g., concerning the use of a particular expression or the structure of a paragraph etc. So human competence for text production includes, at least to some extent, also a kind of *reflective competence*, which machines seem to lack.

Apart from these critical observations there is at least one interesting lesson that AI-generated content could teach us, a lesson which I already suggested earlier on. The fact that many observers commented on the coherence and "fluency" of GPT-3's texts indicates that for human readers certain surface signals play an important role in their assessment of *local* discourse coherence and that the feeling of a "smooth" flow of text contributes essentially to the impression of high text quality that many observers experienced. Whereas, on the other hand, the feeling of a certain "shallowness" of some texts could be attributed to the relative standardisation of these texts that is caused by the model's preferences for statistically highly frequent linguistic material. Finally, *global coherence* still seems to be a weak point of GPT-3 and its relatives, a weakness that shows up in problems like repetitions, unmotivated topic shifts and other problems of subtle topic management. Generally speaking, GPT-3 still needs a little help from its human friends, both by devising the right kinds of prompt and by comprehending the final product. In spite of its remarkable achievements, GPT-3 is still far away from Artificial General Intelligence (AGI).

12. Coherence in dialogue

12.1 What is coherent dialogue?

After discussing the basic resources for the achievement of coherent discourse and some of the applications of this approach to various text types, to multimodal discourse, and to questions of hypertext and intertextuality, I now turn to some of the features of dialogue that make establishing coherence in dialogue a special task and to the practice of using the fundamental resources of discourse in conversational contexts.

Basically, all the organising principles of discourse that I discussed so far play an important role for the coherence of dialogue as well, e.g., common ground and its management, speech-act sequencing patterns, context-change monitoring, reference and coreference, and topic management. However, they are applied in some very special ways, e.g., in the joint construction of common ground, with speech-act sequencing interacting with turn-taking, in the dynamic co-construction of emerging topics, and in the use of specific linguistic means of spoken language.

Starting my exploration, I will now first look at some of the features that make the consideration of coherent dialogue special. In a general, vague sense we seem to know quite well what "coherent dialogue" is. Metaphorically speaking it is the smooth, orderly flow of conversation that is sometimes taken to be the prototypical manifestation of coherence in dialogue. Now this should not be taken in too narrow a sense, as in naturally occurring talk many things happen that do not make for a superficially "smooth" course of dialogue, e.g. interruptions, corrections, afterthoughts, clarifications of misunderstandings, changes of topic, and digressions. All these events, however, belong to the normal organisation of talk and should not generally be assumed to cause lack of coherence.

Furthermore, there are cases of unfriendliness or impoliteness which may be disagreeable but not incoherent in principle. Unfriendly and impolite acts can make sense in certain contexts and types of communication and can be considered coherent contributions in such contexts. This is true, for instance, of the ritual insults traded by boys in South Central Harlem I mentioned in the chapter on discourse understanding (chapter 9).

On the other hand, an extreme accumulation of non-cooperative acts may prevent the participants from fulfilling their communicative tasks and cause disruption to the degree that meaningful interaction is not possible. As an example, I quote from a dialogue written by the famous German cabaret

performer Karl Valentin, where one of the participants asks the other all kinds of questions about his personal life and this person always answers by saying "no":[1]

(1) V: Do you have children?
B: No.
V: How many?
B: No.

In his second question ("How many?"), V violates a presupposition rule for this type of question – it presupposes that it is common ground that B has children – and B, in their answer, violates a rule for *answers* to *how many*-questions. What is interesting about this dialogue is that there is (at least) one interpretation that could make sense of these rule violations: V cannot stop asking his stupid questions, and B is fed up with this interrogation and refuses to give sensible answers.

But this dialogue obviously indicates one of the limits of coherence, as producing a coherent dialogue is, in a sense, always a collaborative effort. But between prototypical "smooth flow" and pure disruption there are many shades of grey.

Taking into account these reflections, a useful strategy to address the issue of coherence in dialogue could consist in showing how interactants step by step produce "meaningful successions or 'sequences' of actions or 'moves'" (Schegloff 2007, 2), using various resources like common ground, sequencing patterns, topic structures and turn-taking rules, and "to see how, even when misunderstandings and trouble arise, these can be coherently shaped by sequence structure in conversation" (Schegloff 1990, abstract). This is the strategy I shall pursue in the following sections.

12.2 Turn-taking

The most basic characteristic of dialogue is turn taking. It is therefore not surprising that the organisation of turn-taking also plays a central role for the coherence of dialogue.[2]

In its simplest form turn-taking takes place in a speech exchange of two persons face to face, with joint attention and (a considerable amount of)

---

[1] Karl Valentin (1969, 169; my translation).

[2] A classic analysis of turn-taking is the paper by Sacks, Schegloff, and Jefferson (1974). Useful contributions I shall draw on in this section include Ford/ Thompson (1996) and Schegloff (2007).

common ground. The flow of the dialogue is organised by (regular) speaker change, and with each turn each participant performs a speech act (or a series of acts), in some way addressing the speech act (or acts) performed by the other participant in the preceding or following turn. A simple example is a question-answer sequence of the following type:[3]

(2) A: Where did they go?
B: They went out to dinner.

In this simple two-part sequence, a great deal of discourse "apparatus" is deployed, which I shall briefly indicate here. The participants follow (a variant of) the speech act sequencing pattern "question & answer", using standard syntactic forms for their respective acts. By asking B a question A selects B as the next speaker, so the allocation of the second turn is determined by the kind of act A performs with the first utterance. By performing the question initiating the sequence A incurs the commitment that A does not know, where "they" went and that A would like to know it. A and B, in using the pronoun *they* to refer to persons known to them, make use of their common ground, and they also build up additional common knowledge in this exchange: A did originally not know where "they" went, whereas B knew. Now they both know and know that they know. Thus, we now have a changed context. Potentially, this sequence could open up a (new) topic, e.g., the topic of the life "they" are leading at present – which in the original conversation is not the case. So what we see here in its simplest form is that turn taking interacts in subtle ways with other aspects of the coherence of dialogue. What is interesting about the timing of second utterances is that they frequently occur with only a minimal gap, so that the speakers apply strategies of "precision timing", a phenomenon we shall come across several times.[4]

One of the reasons why the transition from one speaker to the next frequently functions so smoothly is the fact that in many turns there are "transition relevant points" (TRPs) which are characterised by the completion of a speech act, indicated by means of syntax and also intonation.[5] In some cases they can also be indicated by gaze and gestures. By perceiving such TRPs and also by anticipating such points, speaker/hearers can recognise and "predict" a possible end of turn and fit in their own turn. A simple example given

[3] This short bit of dialogue is extracted and slightly adapted from a transcription in the Santa Barbara Corpus of Spoken American English (SBC050).

[4] For the use of the term "precision timing", see Jefferson (1973). On the remarkable speed of speech-act recognition, see Levinson (2016, 1).

[5] On the combination of action features with syntactic and intonation features in complex TRPs, see Ford/Thompson (1996).

by Ford and Thompson, which shows how action pattern, syntax and intonation coincide, is the following:[6]

(3) C: They never drink/ without it /. >
D: Yeah/ you never have liquor/ without (1.0) fried meat/. >

The first speaker's TRPs does not *have* to be used by the interlocutor to start their own turn. However, it provides an *opportunity* to do so. In many cases there are reasons for the interlocutor to assume that the first speaker will continue their turn with additional acts, for instance in order to support the first statement or to clarify the topic they want to introduce. And, as Ford and Thompson found in their data, a considerable number of speaker changes were initiated at non-TRPs (Ford/Thompson 1996, 157). Functionally, these speaker changes were often associated with the display of affiliation or disagreement with an ongoing turn. Some of the examples in the following will be of this kind.

Not in all cases of speaker change we have the above-mentioned (small) gap between the utterances of the two participants. A different type of change is the following variant of example (1):

(4) A: Where did they [go?]
B: [No idea]

Here the brackets indicate that we have an overlap of utterances by the two speakers. B anticipates the purpose of A's question and starts speaking before A actually finishes the utterance, creating the overlap. This is again a remarkable feat of timing. With this practice B could maybe signal early on that there is no point in asking the question as B does not know the answer. Such quite frequent "early starts" indicate that participants, with a rich background of common knowledge, can frequently anticipate the other's intention quite early on and perform a coherent response when the original speaker is still finishing their utterance. In some cases, this practice can be used to signal additional information, e.g., that the second speaker is impatient.[7]

An interesting version of "early response" occurs in the context of practical activities like cooking together or repairing something in the household together. In such contexts, a request made by person A, e.g., the request to

---

[6] Ford/Thompson (1996, 157). In the transcription, the slash indicates a syntactic completion point, the full stop indicates a "final" intonation contour, and ">" indicates the completion of an action, in the case of Cs turn his statement. Underlining indicates emphasis.

[7] It is important to remember that overlapping talk is a surface phenomenon that can be used either in a "cooperative and rapport-building" way or as an "interruptive" device (see Tannen 2007, 94).

hand A a screw-driver, will frequently be complied with by person B before the request has actually been fully expressed.[8] In these cases, it is the combination of various resources that facilitates the early (practical) response: knowledge of the practical routine, joint attention to the activity at hand, knowledge of the sequencing pattern "request$_A$ & performing the requested action$_B$", and the knowledge of the linguistic form of typical requests that allows recognizing an intended request early on.

In other cases, short interruptions are used to insert corrections or other kinds of "repair", as in the following example:

(5) A: In the forest we saw a raven. [It made strange] noises. OK, a crow.
B: [a crow]

Here the amateur ornithologist B corrects the other's bird identification and A accepts the correction. This kind of correction is often found in narratives where two persons experienced certain events together and one person takes the leading role in narrating the events and the other contributes corrections, additions etc.

A remarkable variant of this type consists in short insertions of B in A's flow of speech which are intended to help A to find the correct word or syntactic construction. I remember a conversation with an aphasic patient who had a serious word-finding problem and whose wife had perfected the strategy of supplying a missing word at the very point where her husband should have used it in his flow of speech.[9]

Starting one's own turn before the other participant has finished their turn can also have various other functions. One well-known function consists in "interrupting" speaker A, which can serve to signal that B considers A's contribution to be irrelevant and/or to introduce a new topic preferred by B. If this move is successful, it creates a topic boundary, which is interesting for an assessment of the coherence of this stretch of dialogue. In other cases, the main function of interrupting and talking in parallel with another participant consists in "gaining the floor" through a rather rough technique of "self-selected" turn-taking. In many cases this is also connected with the interruption of the ongoing speech act sequence and/or topic. If the active speaker does not give up their turn voluntarily, this can lead to an extended phase of overlap, what we call "fighting for the floor".

---

8 For a detailed analysis of precision timing in such sequences, see Deppermann/Schmidt (2021).

9 Some aphasia therapists discourage partners from using this remarkable strategy, because they feel it tends to lessen the patient's motivation to improve their word-finding performance.

From the point of view turn-taking – and of course in respect of the act structure – story-telling is a very special pattern. In order to prepare the interlocutors for an extended turn and to gain the floor the prospective narrator frequently declares their intention by making a preparatory move like the following:

(6) What I wanted to tell you …

(7) Something strange happened to me today.

(8) You wouldn't believe what I saw today.

Moves like (7) and (8) not only serve to rouse the listeners' interest but also to signal the kind of story, e.g., a strange and incredible story. This kind of signal also indicates the kind of reactions that the narrator will assume as appropriate, namely expressions of astonishment and incredulity.

Like many other types of sequence, story-tellings usually invite so-called "backchannel" contributions, i.e., short utterances of amazement, empathy or assessment, which do not deserve a turn of their own but which serve as "keep-going signals" and contribute to helping the story-teller along.

In the joint story-telling situation mentioned before we sometimes also find a very special type of turn-taking strategy, namely the strategy of telling the story taking turns. This option is, for instance, used when two persons travelled together and now recount their shared adventures. In this case the coherence of the story is produced as a cooperative effort.[10] This does, however, not preclude competitive elements, as everyone knows who has listened to couples telling such stories.

As I mentioned at the beginning of this section, turn-taking is always linked to other organising principles of dialogue, e.g., speech act sequencing, topic management, and knowledge management. I shall now leave these brief observations on turn-taking and move on to other aspects of dialogue organisation, starting with common ground management. On various occasions this will provide the opportunity to return to matters of turn taking.

[10] For an example of a "polyphonic" style of collaborative narrative, see Chafe 1997.

12.3 Managing common ground in dialogue

Like in the production and reception of texts, common ground plays a fundamental role both as a resource and a result of dialogue. Take the following short extract from a conversation:

(9) DANA: So I don't know what they're doing. But I can call, and find out. If they wanna come over.
NANCY: .. Okay.

To understand this short exchange, we have to know at least who *they* refers to and what *coming over* implies – e.g., from somewhere in the same building or from next door.[11] To see the point of Dana's utterance we have to know, in addition, that the participants are two friends (students) planning what to do in the evening. Having taken part in the earlier part of this conversation, the participants obviously know this and this shared knowledge makes it possible for them to interact in this way. To get a more subtle understanding of this extract we would probably also have to know why it is Dana who makes this suggestion and what role this suggestion plays in a sequence of earlier suggestions etc. All this knowledge can be assumed to be available to the participants of this conversation.

While presupposing and using certain bits of pre-existing knowledge, this exchange also *produces* extra shared knowledge for the parties to the conversation. After the exchange both persons know that Dana does not know what *they* are doing (at the moment) and that she is willing to call and to find out if they want to come and join the girls. Part of the relevant knowledge produced by an utterance concerns the commitments a speaker undertakes by making a certain speech act, for instance by referring to a person or by making a suggestion as in this case. For this knowledge of the two persons to become common ground for them, Dana has to assume that Nancy knows this, and Nancy has to assume that Dana knows it. Sharing this body of knowledge (and assumptions) as common ground, the two speakers can use it as a resource in the further development of the dialogue.[12] Nancy could, for example, complain later on that Dana did not call although she herself suggested doing this etc.

As in this example, common ground can often be unquestionably assumed to be available to the participants, and its incremental growth will also be

[11] This conversation is discussed in more detail in the following chapter.

[12] This presupposes, as discussed in chapter 3 of this book, that the participants of a conversation keep track of everybody's commitments and the perspectival character of this structure.

unproblematic. This is partly due to fact that speakers routinely take up the cues to assumptions provided by the utterances. For instance, in our above example the use of the pronoun *they* (in this context) implicitly signals that these persons are assumed to exist and to be known to the participants. Similarly, the use of definite descriptions like *the old tree* or *the Covid death toll* are mostly used with the presupposition that the entity is known. And Dana's making her suggestion signals her assumption that the suggestion can be carried out and that she is prepared to execute it. Thus, it is by means like these that parts of common ground are implicitly managed.

In many cases, however, it becomes necessary to make sure that there is in fact enough common ground and if not, to create it. For this purpose, various devices are available.

A first device, which is frequently used at the beginning of narratives, consists in giving some information on persons, places and other background elements that might become relevant in the course of the narrative as in the following example:[13]

(10) At the time, I was living in Washington, D.C., working as a paralegal for a high profile law firm which specialized in aircraft crashes. Yes, before I got this job, I had not realized that planes crashed often enough so that a whole law firm could be devoted to only handling such matters.

This section of a story has been called the *orientation* section by Labov and others.[14] It is, however important to note that, depending on the kind of story, this type of information can also be given spread throughout the narrative, provided "just in time" at the point where the knowledge is needed.

Probably the most fundamental device for securing common ground in dialogue are *question-answer sequences*. These can occur in various forms and contexts.[15]

A first type of question-answer sequence can be used at the beginning of a stretch of conversation where speaker A tries to make sure that speaker B possesses knowledge relevant to what A wants to tell. For instance, if A wants to tell B something about their research on de Bruijn graphs, they might ask B:

(11) A: Do you know de Bruijn graphs?
B: Yes, I studied them some time ago.
A: OK. In my project …

---

[13] https://medium.com/afrosapiophile/the-death-of-a-mother-6ec0702515c6. (Accessed 18.10.21).

[14] See Labov/Waletzky (1967).

[15] For further aspects of question-answer sequences, see the following chapter.

If B does not answer in the affirmative, A can then decide to either give a short explanation of what these graphs are like or to decide not to tell his story at all.

A well-known version of this type of sequence occurs in direction-giving dialogues. Let us assume that after A's inviting B to her house B asks how to get to A's house. The dialogue might run as follows:

(12) B: How do I find your house?
A: You know our village?
B: Yes, I was there once. I remember the church.
A: OK. You drive uphill past the church and then …

Before answering B's question of how to find A's house A inserts a question, the point of which is to find out what knowledge B has that A can use in her direction-giving. After B answers the question by telling A that she remembers the location of the church, A can use this bit of knowledge in explaining B where to go. So what we have here is a question-answer sequence inserted into the direction-giving for the purpose of securing relevant knowledge.[16]

Whereas in this case the "active" speaker, i.e., the one giving the directions, initiates the question-answer sequence, in other cases it is the person in the (temporary) listener role who inserts a question for *clarification*, thereby initiating a question-answer sequence. In the following example the question concerns the meaning of an expression the first speaker used:[17]

(13) A: But the air's gotta come in there and the air is sorta infiltrated with little uh pixy dust.
B: *Pixi* dust?
A: You know from the big boom.
B: *Pixi* dust?
C: "Radioactivity" I think is what he means.
B: Oh. Okay.

A is talking about a bomb shelter and jocularly uses the expression *pixi dust* to refer to radioactivity. B asks for the meaning of the expression, and in his explanation, A uses another jocular euphemistic expression (*the big boom*), and B has to repeat the question. Finally, C helps by giving a more adequate explanation (*radioactivity*), and B accepts the explanation, so that A can continue with his story. In this case we have a slightly extended question-answer sequence initiated by B, which serves to secure shared knowledge about the meaning of A's expression *pixi dust*.

---

[16] On "insertion sequences" of this and related types, see Schegloff (1972).
[17] For a full transcription of this conversation, see Jefferson (1972, 336).

Clarification sequences of this type, which can be an indicator of misjudged common ground on the part of the speaker, may relate to various aspects of the meaning of parts of a speaker's utterances, e.g., to the meaning of special terms, complicated expressions, and also to the implicit assumptions of the speakers like in the case of irony and other kinds of "indirect speech acts". Thus, they are an important repair device for maintaining shared knowledge and the coherence of the conversation.

Before leaving this topic, I should like to mention a category of quite inconspicuous devices that also contribute to the securing of shared knowledge, namely so-called *backchannels* like *mhm* or *uh huh*. These backchannel elements are quite multifunctional, as they can indicate that the listener is still attentive, that they understand what is being said and that they are monitoring the common ground development. Take the following two examples:[18]

(14) A: and their lawn is high up, up the road, you know
B: Mm
C: Yeah
A: so he dug up the lawn and put a little garage in

(15) L: Yeah, you just got to be caref- well see Dwight only has one gall bladder?
E: Mh hm
L: He had- and then he has to be careful what he eats, he can't eat anything greasy or anything, you know
E: Mh hm

In both cases the backchannels are placed at the end of speech acts presenting small information units, and in this position indicate that the hearer is following the speaker and taking up the information given.

What this brief section shows is that participants in a conversation have at their command quite a wide repertoire of means with which to secure common ground and thereby to contribute to the coherence of the conversation.

Before leaving this section I, have to mention an important aspect of the dynamics of common ground that so far only briefly figured in my account. The parties to a dialogue will sometimes be of different opinion and this disagreement can be made explicit by contradicting another party's assertion(s). In such a case the disagreement will be mirrored in the commitment-store of the participants as follows:

[18] Biber et al. (1999, 1091); Jefferson (1984, 199).

(16) (a) A is committed to the belief that p, (b) and B is committed to the belief that not-p. (c) And A knows that B is committed to the belief that not-p, (d) and B knows that A is committed to the belief that p. (e) And (16c and d) is common knowledge of A and B.

If for some reason A becomes convinced of B's position, they can retract their original assertion, a move that changes A's set of commitments. But from the point of view of the "history" of this dialogue, both the original set of commitments and the retraction remain in the history of common ground, so that A or B may later refer to this change of mind, if necessary.[19] Of course, there are much more complicated pathways of dialogue, where an agreement takes long to reach or never comes to pass. In any case, the coherence of such a course of dialogue is based (partly) on sequencing patterns of the following or a similar kind: "assertion$_A$ & contradiction$_B$ & arguments (for the contradiction)$_B$ & retraction$_A$ & accepting retraction$_B$" etc.

12.4 The language of dialogue

A point that has surfaced on several occasions in this chapter is that in dialogue/conversation speakers use linguistic resources that are in some ways different from what is usual in written texts.[20] Among these resources we find, for instance, short (condensed or elliptical) versions of sentences, "lexical bundles" like *I don't know why*, discourse markers like *well*, *you know*, and *right*, interjections like *Oh* and *wow*, exclamations like *How cool!*, deictic expressions like *over there!*, response forms like *yes* and *no*, so-called utterance openers like *I'll tell you what*, tags like *isn't it?*, *or something* or *not at all*, and the backchannels mentioned before.

Generally speaking, utterance types used for the performance of speech acts and speech act sequences are more flexible than the standard sentence types used in written texts. The reason for this can be found in the typical situation in which spoken language is produced: "We may refer to the grammar of speech as 'dynamic', in the sense that it is constructed and interpreted under real-time pressure" (Biber et al. 1999, 1066). One of the (historical)

[19] On retractions and corrections in dialogue games, see Hamblin (1970, 263-268), Asher/Gillies (2003).

[20] A useful survey of the "grammar of conversation" of English is provided in chapter 14 of the Longman Grammar (1999). In the following, I shall use some of the terms and examples given there. For a survey of some basic aspects of spoken German, see Zifonun/Hoffmann/Strecker (1997), vol. 1, chapter C.

consequences of this situation is the emergence of a grammar of short utterance types, which appear condensed or elliptical as compared to standard written sentences, e.g. *More sauce?* (vs. *Would you like some more sauce?*), *Remember?* (vs. *Do you remember?)* or *Not at all!* (as an answer to *You don't mind?)*. Another consequence is that we frequently find an "add-on strategy" (Biber et al. 1999, 1068) for the production of utterances, including afterthoughts, clarifications, corrections, and tags of various types, instead of the complex hierarchical structures we tend to find in the grammar of written texts. One interesting example of such a strategy is the use of right dislocation for the clarification of reference, e.g. *I never met him, your brother*.

This brief section on the language of dialogue is, of course, only an elementary hint at the fascinating topic of spoken language, which will be surfacing in most of the examples to be discussed in the following chapters.

From here we now go on to the question of sequencing patterns in dialogue, which in some respects are different from monological patterns but in others show clear similarities.

13. Speech act sequencing in dialogue

In dialogue, speech act sequences occur in two types of arrangement, sequences *within* turns and sequences *distributed over two or more turns*. Sequences within turns are frequently quite similar to the monological sequences I discussed in chapter 3, whereas sequences distributed over moves are typical of dialogue. In this chapter, I shall therefore concentrate on discussing the latter type.

What this chapter is intended to show is, among other things, that sequencing patterns play a fundamental role in the creation and monitoring of coherence in dialogue, regularly interacting with other factors like common ground and topic continuity

13.1 Question-answer exchanges

In many cases, sequencing of speech acts and turn taking in dialogue go in step, i.e., a sequence pattern is realized in two turns. Examples of such sequences, which could be called two-turn moves, are "question & answer", "greeting & greeting", "complaint & apology", "compliment & counter-compliment" etc. In studies on Conversation Analysis such patterns have been called "adjacency pairs".[1]

Basic features of such sequencing patterns are the following:

(i) Speaker A initiates a (stretch of) dialogue by uttering the first member of such a sequence (act_1).
(ii) Act_1 places a "conversational demand" on speaker B to respond to A.
(iii) Act_1 opens up a space for response, including a (varying) array of specific replies (act_2).

I shall now go into some of the ramifications of (iii) by discussing some examples.

Probably the prototype of this kind of sequence is the "question & answer" sequence. Some of the typical manifestations of this pattern are shown by the following example:

---

[1] For the concept of adjacency pair, see. Sacks 1992, vol. II, Part VIII (Lectures Spring 1972). The functional motivation of sequential organisation (e.g., adjacency pairs and their extensions) is discussed in Kendrick et al. (2020). See also section 4.1 in the present book.

(1) A: What are you going to do tonight?

B: (i) I'll be staying at home.
(ii) I don't know yet.
(iii) I'm not telling you.
(iv) You don't really want to know.
(v) Why do you ask?
(vi) It's none of your business.

All these replies by B are relevant to A's question, although not all of them are cooperative in the same way. They indicate in what way B understands the kind of act performed by A and the "conversational demand" set up by this kind of act.

B (i) is what one could call a "straight answer", maybe the kind of answer A expected. But the others also address aspects of the question, mainly commitments A undertaken in asking the question. These commitments could be formulated as follows:[2]

In asking (1) A is committed to the assumption that:

(ii) B knows what B is going to do.
(iii) B is willing to tell A what B is going to do.
(iv) A wants to know what B is going to do.
(v) A has a reason to ask what B is going to do.
(vi) A has a right to ask what B is going to do.

Not all these commitments are strictly necessary for A's meaning and B's understanding A's utterance as a question for information, and there are others, which are related to different functions of the question. If A makes the additional assumption that B should stay at home and study and B knows this, B could understand the utterance as a *reminder* that B should study. Or if B assumes that A wants to go out with B tonight, B could understand A's utterance as a preparatory move for an invitation to go out.

As for the relevance of the replies in (1) to the question asked and the coherence of the resulting sequences in (1), we can now assume that it is the special relation of the replies in (1) to the commitments of the question that accounts for the relevance and coherence. Thus, although B's (1)(iii) is a refusal to answer the question and (1)(vi) rejects the questioner right to ask the question, these utterances can be accepted as coherent contributions to the exchange.[3]

---

[2] The number of the commitment refers to the respective reply in (1).

[3] In the case of (1)(iv) (*You don't really want to know*) there is also a special use indicating that A will be doing something that A will not like.

In the case of a further meaning of the question, for instance the reminder version with its additional commitments, other replies could be explained as coherent that may appear *prima facie* incoherent, e.g., the following:

(2) A: What are you going to do tonight?
B: Oh, I know, I know.

Or in the case of the pre-invitation:

(3) A: What are you going to do tonight?
B: (i) I have a headache.
(ii) I can't go out. I have to study.

Both utterances of B can be understood as rejections of the projected invitation. Understanding B (i) may need a – not very difficult – inference on the part of A, and in B (ii) the additional explanation provided by B can serve to mitigate the rejection of the projected invitation.

In addition to these basically cooperative replies there are also other possible responses that refer to aspects of the common ground situation:

(4) A: What are you going to do tonight?
B: (i) Leave me alone, I have to concentrate.
(ii) You always ask me the same thing.

Generally speaking, the gamut of possible responses to a question that will be counted as coherent is considerable, and if the participants have reason to assume a basically cooperative attitude of their interlocutors, the presumption of coherence will make many utterances comprehensible and will instigate inferences in the case of utterances that may not be so at first sight. Of course, there are "question & answer" constellations in which both possible questions and coherent answers are much more restricted, for instance in the case of examinations and interviews.

The main cause of apparently incoherent responses, as I mentioned earlier on, is misunderstanding. Therefore, cooperative interlocutors will generally see apparent incoherence as an indicator of misunderstanding and try to locate the cause. This kind of thing sometimes occurs in conversations with children. The following example is taken from the book by David and Rosa Katz "Gespräche mit Kindern" ("Conversations with children"). The exchange begins with a role play of the two boys of the family (4 and 5 years old) where the (younger) boy says "I am the baby" and his mother, who is also present, replies "You will probably always be the baby". She then suggests to the elder brother that they should begin to use their proper given names, and he replies that he wants to continue being called his nickname,

following which she asks the younger boy if in future he wants to be called by his given name(s) (*Julius Gregor*) instead of his nickname *Baby*:[4]

(5) Mother: Baby, shall we call you Julius Gregor or Julius or Gregor?
Boy: I want to be Daddy and have a motorbike.

In their commentary, the authors explain that the boy obviously did not understand the point of the question. Seeing that he has so many options to be someone else than Baby, the boy decides to choose an extra option, namely to be Daddy and have a motorbike. His response could also be interpreted as a continuation of the earlier role play, which interpretation also presupposes a misunderstanding of the question. Taken as the expression of a choice of role in the continuing role-playing exchange, the utterance would make sense, even though it appears "locally" incoherent.

Apart from *bona fide* misunderstandings there are also cases of jocular non-cooperation like the intentional misinterpretation of idiomatic uses of sentences which also have a "literal" reading like in the case of the following request:

(6) A: Can I ask you to help us?
B: Yes, you can.

Assuming that A's utterance in (6) was meant as a polite request for help – for which *Can I ask you* is a standard formula – and that B understood it this way, B's response is a case of feigning a misunderstanding, seemingly interpreting the utterance as a question concerning A's ability or his being allowed to something.

Finally, we have the practice of evading a question, which also leads to problematic question-answer sequences. Evading a question is generally considered a favourite ploy of politicians. The following example is taken from a newspaper report on a meeting of the liaison committee of the British Parliament (07.07.2021):[5]

(7) Of all the MPs questioning, Chris Bryant probably received the least satisfaction in the moment […]. "Did you sack Matt Hancock?" he asked [the prime minister] bluntly, and was rewarded with a statement so evasive it sounded like a parlour game, a cross between Just a Minute and Truth or Dare. "We read about the story concerning Mr Hancock and CCTV on the Friday and we

[4] Katz/Katz (1928, 46; my translation).

[5] https://www.theguardian.com/politics/2021/jul/07/boris-johnson-performs-another-masterclass-in-evading-mps-questions. (07.07.2021) (The health secretary Matt Hancock had announced his resignation from government two weeks earlier after admitting breaking Covid rules.) On another example of an evasive response the journalist commented: "irrelevant to the point of senselessness".

had a new health secretary on the Saturday. That's all I have to say on the matter."

Thus, what could look like a misunderstanding of the point of the question causing an incoherent question-answer sequence is interpreted by the journalist, probably rightly so, as a case of intentional evasion on the part of the prime minister.

13.2 Extensions of question-answer sequences

Generally speaking, most two-part sequences can be extended to form longer sequences. In the case of question-answer sequences, there are various types of extension, of which I shall mention two here.

A first type of extension consists in various types of explanation connected to either the question or the answer or both. One can, for instance, explain why one asks the question, and the answerer can explain why they give this particular answer.

(8) A: (a) Excuse me, I have a problem. (b) Could you tell me how this machine works?
B: (a) I'm sorry, I don't know it myself. (b) I never use it.

In (8) A prepares the question by giving a reason why they ask it (a), whereas B gives a reason why they have to give a disappointing answer (b). Explanations of the latter type can also serve to indicate that the answerer is not just being uncooperative.

A second type of extension consists in creating common ground for the question and/or the answer as a kind of local knowledge management. Take the following example from "Oral Answers to Questions" in the House of Commons:[6]

(9) **Sally-Ann Hart** (Hastings and Rye) (Con): [903092] (a) In many villages, such as Winchelsea Beach in my beautiful constituency, sewage or foulwater drainage is not separate from rain water and storm water drainage. (b) There is a combined system, which means that at times of high rainfall, many residents suffer sewage flooding in their gardens. (c) Does my right hon. Friend agree that in view of environmental and health concerns, it is imperative that sewage is kept in a separate pipe from rainwater? (d) If so, what steps will he take to ensure that these changes happen?

[6] Hansard, *Parliamentary Debates* Volume 699, No. 39, p. 1119 (22 July 2021); the answer is given by The Secretary of State for Environment, Food and Rural Affairs (George Eustice).

(10) **George Eustice:** (a) My hon. Friend makes a very important point. (b) Some of the challenges we have are typically with houses built in the Victorian era where, as she says, the street drainage system goes into the foul water sewage system. (c) That can lead to it being overwhelmed at times. (d) Most developments that have taken place since the 1960s do have surface water drainage separated from foul water sewage systems. (e) We have set up a taskforce to look at how we can address this problem and, in particular, reduce the use of combined sewage overflows.

In this exchange we have two related questions put by the member of parliament and answers to these questions, which are both surrounded by extra statements. In (9) the first two statements (9a) and (9b) serve to describe the problem that occasions the questions, thereby creating common ground for the questions. In (10), the secretary of state first politely acknowledges the importance of the question put by a member of his own party and then goes on to describe the problem himself, specifying it ("houses built in the Victorian era") and restricting its relevance mainly to developments before 1960. Finally, in (10e), he answers the main question (9d) by stating what has already been done ("set up a task force") and what the assignment of the task force is. With his repetition of the original problem description, which explicitly refers to this description ("as she says"), the secretary of state can document his recognition of the problem, albeit slightly downplaying it, before actually answering the question.

The upshot of these observations is that the two elements of a two-part sequence like "question & answer" can be surrounded by quite a number of acts related to them and that, therefore, the distance between the two elements can be quite considerable without affecting the coherence of the dialogue, as other examples in this chapter will show.[7]

13.3 Making an accusation and answering the accusation

A somewhat more complicated sequence pattern is "making an accusation & answering the accusation".[8] "Blaming (someone for doing something)", "complaining about an action", and "criticising an action" with their relevant responses are closely related patterns, however with interesting individual features and corresponding response patterns. Generally speaking, in every-

[7] For a detailed analysis of an adjacency pair of the type "request & granting of request" and its extensions, see Schegloff (1999).

[8] For similar patterns in written controversies, see Fritz (2005). An example from the Bramhall-Hobbes controversy was given in chapter 3.

day life accusations are frequently used to discipline persons (e.g., children, pupils) and to inculcate norms.

As I already showed in chapter 2, we can describe some of the features of accusations by noting the following aspects:

A can accuse B of having x-ed by asserting that B x-ed by uttering "B x-ed".

In doing so, A undertakes the following commitments:
(i) A believes that B has in fact x-ed,
(ii) A assumes that A can provide evidence for B's x-ing,
(iii) A assumes that B should not have x-ed and that B knows this,
(iv) A assumes that B is responsible and could have acted otherwise.

In the earlier chapter I tried to show how this inventory of commitments (and others) can be used by the speakers as a resource for their following acts in the sequence, e.g., by making it explicit that they can provide evidence for B's deed.

Generally speaking, many aspects of actions can become the target of accusations. One can, for instance, accuse someone of having acted too slowly or too impulsively. In some cases, the aspect of intentionality, which is generally presupposed for actions, can be particularly emphasised by adding that the defendant acted deliberately (as opposed to clumsily). This may then be denied by the defendant. Similarly, the aspect of motive may be emphasised ("he acted out of base motives"), which again may become the target of a rebuttal.

What I shall do now is point out how the batch of commitments of A mentioned above also contributes to setting up a conversational demand for B and, at the same time, can be used by B as a resource for relevant replies.[9] A group of such replies is the following:[10]

(11) A: You stayed away last night.
B: (a) I am sorry. I should have come back.
(i) That is not true, I was back home at midnight.
(ii) How do you know? You were sleeping.
(iii) So what? I was allowed to stay away.
(iv) I couldn't get a bus back. So I had to stay at X.s

[9] Two early analyses along these lines are Fritz/Hundsnurscher (1975) and Muckenhaupt (1978). Many useful observations on "remedial interchanges" and excuses are to be found in Goffman (1971, ch. 4) and Austin's "A plea for excuses" (Austin 1970, ch. 8).

[10] The numbers of the replies refer to the numbers of the commitments mentioned before, with the exception of the apology B (a), which has a special status.

The first reply in this list, B (a), an apology, has a very special status. This status was appositely characterised by Goffman in his "Relations in public" as follows (Goffman 1971, 143):

> In its fullest form, the apology has several elements: expression of embarrassment and chagrin; clarification that one knows what conduct had been expected and sympathizes with the application of negative sanction; verbal rejection, repudiation, and disavowal of the wrong way of behaving along with vilification of the self that so behaved; espousal of the right way and avowal henceforth to pursue that course; performance of penance and the volunteering of restitution.

Depending on the situation and the kind of misdeed or misbehaviour not all these aspects are relevant in the same way, but this description represents quite well the "aura" of a simple apology like the one in our example and shows how it could be extended. In this case B simply utters an expression of regret, thereby accepting the presupposition that he/she had in fact x-ed, and clarifies that he/she "knows what conduct had been expected". In many cases this ritual act is sufficient to remedy the situation.

The second reply, B (i), attacks the assumption that B in fact did what he/she is accused of and supports this attack by telling what *really* happened. This kind of defence can sometimes be massively extended by providing a narrative of the real events as seen by B. This type of move may lead to a discussion of matters of *proof*.

The third reply, B (ii), attacks the assumption that A has evidence for B's x-ing and, in addition, gives a reason why A cannot have this evidence. Again, this support of B's attack can be extended by giving more detail of the problematic situation.

The fourth reply, B (iii), refers to the norm presupposed by A's accusation and states that it does not apply in this case. In this defence B "accepts responsibility but denies that it was bad" (Austin 1970, 176).

Finally, with the fifth reply, B (iv), B denies responsibility, thereby possibly accepting the view that it was bad. Again, this move, a kind of justification, could be very much extended by telling in detail why it was impossible to catch the bus, for instance in a case of *force majeure*.

This brief description of a core of the sequencing pattern of "accusing & answering an accusation" already shows some of the points where the second move opens up further possible courses of the exchange: If the accused denies the fact, this opens up the space for a *debate* on what the facts are. If the accused denies the relevance or applicability of the norm, this can lead to a debate on the norm and is application. And finally, if the accused uses the justification move, the question of their responsibility can become a matter of debate. In any case, the details of the problematic event may have to be de-

scribed and discussed in more detail. Some of these further ramifications are also prefigured in the third moves, e.g., in the move supporting the straight denial ("… I was back at midnight.").

Thus, what we see here is that two-part sequences have the potential for opening up a wide space of moves related to their basic structure. In some cases, this space of moves is determined by topics opened up by the second (and third) moves, e.g., the topic of the respective norm or the topic of responsibility. Another point worth mentioning in this context is that the structure described here is only the hard core, so to speak, whereas closer analysis shows that there can be many shades of "extenuation", "palliation", and "mitigation", terms used by Austin (Austin 1970, 177), which could be worth describing in detail.

To give this analysis more colour I will now discuss a real-life exchange where accusations and possible replies play a structuring role, contributing to the coherence of the dialogue.[11] I begin with a brief summary, which is provided with the transcription:

> (The recording begins in a car, and moves to the kitchen of a family home. Main participants are three teenage sisters (Sabrina, Kendra, and Marlena), their mother (Kitty), and step-father (Curt). A friend of Sabrina's (Gemini) is also present. The dispute centers around Kitty's belief that Kendra stayed the night at a friend's house without permission, something which Kendra denies having done.)
>
> KENDRA: You're so stupid thinking I spent the night. I came home last night and told you.
> SABRINA: Kendra, ... just let it go.
> KITTY: [You] --
> KENDRA: [No], because she doesn't --
> KITTY: All you [said last night was, Kim] wasn't stay[ing the night].
> KENDRA: [She thinks .. I spent the night in my jeans]
> If I did spend the night, and I was trying to lie, I would give up. .. Cause then I wouldn't care. ... Cause I knew I deserved it. But I didn't spend the night, and I don't deserve this. ... Kim couldn't spend the night. I told you. .. She could, but then af- -- I wanted to stay at the game longer? And her mom wanted to take her home early, and I'm like, "no= let's stay longer". But her mom wouldn't let her? And so she went home and, I was like, "I'll call you

[11] The transcription is taken (slightly simplified) from the Santa Barabara Corpus of Spoken American English (SBC042). https://www.linguistics.ucsb.edu/sites/secure.lsit.ucsb.edu.ling.d7/files/sitefiles/research/SBC/SBC042.trn.
Transcription symbols: square brackets aligned vertically: overlap; "..": minimal pause; "...": pause (0.2 sec. or more); "- -": aborting projected utterance; "=": prosodic lengthening, emphasis; "XXX": uncertain words; "< …>: analyst comment.

when I get home, and then you come over". And she goes, "no just .. spend the night. We'll .. do this some other weekend." "Okay cool", ... cause she had to go bowling in the morning. <Voice of another person: Talk to Melanie's mom.> Her mom would know.
... I'll have Melanie call you. Her [mom call you. her mom wouldn't lie,
KITTY: ... Oh, [right, right, right, Melanie will call me to confirm your lie,]
KENDRA: .. Melanie lies but],
KITTY: you'll get a hold of her first. I know how that works.
KENDRA: No Melanie's mo=m.
KITTY: I thought Melanie didn't have a pho=ne. ... But I'm supposed to call Melanie's mo=m.
KENDRA:[(SIGHS)] I didn't want you to have her phone number.
KITTY: Yeah, another lie. Remember?
KENDRA: That was a long time ago.
KITTY: Now I'm supposed to re- -- And I'm supposed to trust you every .. [but how am I supposed to know when] you're telling the [truth].
KENDRA: [That was a joke. That was a joke.]
KITTY: and I'm supposed to read your mind and know what you're joking about.
KENDRA: .. No= but, it was a joke.
KITTY: You know Kendra, -- I don't know how many times I gotta tell you, once you lie=, ... once you [lie, you lose th- .. complete] trust.
KENDRA: [XXX Are they having a carwash]?
SABRINA: No.

< short exchange about a shop that is about to open >

KITTY: .. You're not going anywhere [for a while].
KENDRA: Go=d,
I didn't spend the night, that's what makes me so m=ad, I'm grounded for nothi=ng. ... I was home all night last night, I came home and told you.
... And even my cuffs and everything's on my bed.
... So stupid. ... Next time if I leave in the morning, I'll wake you up. "Mom Im .. awake now, I'm [leaving]".
KITTY: [Well you woke] me up last ni=ght, to tell me Kim wasn't spending the night, I don't know why you couldn't wake me up in the morning, to tell me that you were going.
KENDRA: .. I was tired this morning. ... So much fun. .. My hair looks like a mess. ... I left it in a ponytail, I didn't even have time to take a shower.
... I think if I spent the night I would've took a shower.

< short exchange about playing on the volleyball team and the question who cut the grass>

KENDRA: ... See if I spent the night Mom, you'd think I'd still be in my uniform?
KITTY: ... I know she's -- J- she- -- See now she's trying to think of ways to cover her tracks. I know what you're doing. ... Seen this before.

< exchange between Curt and Kitty about various household tasks >

KITTY: ... Okay. ... I'm gonna go to the bathroom. .. Me and Kendra got in another fight.
KENDRA: < addressing Curt > Let me tell you right now, I did no=t spend the night. ... This [is stupid].
<…>
.. This isn't fair, I'm being punished for no reason. Fucking stupi=d.

< exchange between Curt and Sabrina about the oil light in her car >

KENDRA: Mom doesn't believe I didn't spend the night,
... and if I did, I would be taking my punishment, and I didn't.
KITTY: Since whe=n has she taken her punishment.
MARLENA: Why are you --
KENDRA: Grounded for a month, < yells> and I didn't do nothing.
MARLENA: <yells> Mo=m, she didn't spend the night.
KITTY: <yells> Hey, stay out of it. ... Stay out of it, you're not involved in this.
KENDRA: ... Don't bother with her, she'll just get angry.
KITTY: <yells> You been caught in a lie Kendra, and that's the whole thing. That's why you're pissed.
CURT: ... It's gonna be another wonderful Saturday.

< The rest of the transcription mainly presents an exchange between Kitty and Sabrina about shopping, food, clothes and other household matters. >

The extract of the transcribed exchange centers around Kitty's accusation against Kendra and Kendra's attempts to defend herself against the accusation. Kendra's first utterance presupposes that Kitty accused her earlier on of staying the night with her friend Melanie without permission ("You're so stupid thinking I spent the night. I came home last night and told you."). So here we have a piece of common ground that provides the background to most of this exchange. Kendra's move has three elements: She attacks her mother by saying that her mother's assumption is stupid, she denies having stayed the night and she backs her denial with the assertion that she told her mother when she returned home. Her sister Sabrina tries to prevent her from intensifying the conflict ("just let it go"), but she refuses to relent. Kitty responds by challenging the assertion that Kendra told her by asserting herself that Kendra had only told her something different, namely that (Kendra's friend) Kim was not staying.

After these moves the conflict has reached the point where matters of proof become relevant. The main burden of proof obviously lies with Kendra to show that she did in fact not stay the night and, as a piece of evidence, to demonstrate that she in fact told her mother. Kitty, on the other hand, would have to prove that Kendra had not told her when she came home.

I shall now look at Kendra's strategies of dealing with this situation. Her

first move consists in a kind of reductio ad absurdum of her mother's assumption ("she thinks I spent the night in my jeans"). Obviously, she does not consider this argument strong enough, because she directly continues with an interesting counterfactual conditional argument that also refers to two other aspects of the situation, namely her mother's assumption that she is trying to lie and the punishment which she insists she does not deserve:

> If I did spend the night, and I was trying to lie, I would give up. .. Cause then I wouldn't care. ... Cause I knew I deserved it. But I didn't spend the night, and I don't deserve this.

She explains the basis of the principle expressed by the counterfactual and insists on her innocence. As a proof of her innocence this move presupposes that she does indeed follow this principle, which may easily be doubted by her mother. It is remarkable that a few minutes later she uses a similar argument, when addressing her stepfather:

> Mom doesn't believe I didn't spend the night,
> ... and if I did, I would be taking my punishment, and I didn't.

Directly after this move Kendra embarks on a narrative explaining why Kim could not stay the night. This does not really improve her situation in the conflict, as this point had already been granted by her mother.

At the end of the narrative passage one of the persons present comes out with a suggestion: "Talk to Melanie's mom." Kendra directly takes up this suggestion, a move that appears quite strong as it could provide evidence to decide the conflict: "Her mom would know." Remarkably, this suggestion is not taken up. Kendra's mother is sidetracked by the immediately following utterance:

> KENDRA: ... I'll have Melanie call you. [Her mom call you. her mom wouldn't lie, .. Melanie lies but], No Melanie's mo=m.
> KITTY: [right, right, right, Melanie will call me to confirm your lie,]

Directly after Kendra's slip of the tongue – saying "Melanie" instead of "Melanie's mom" – and before she can make her correction ("her mom"), which she explicitly repeats ("No Melanie's mom"), Kitty attacks this suggestion, taking up the topic of lying which dominates the next few utterances. What is interesting about this passage is the way how Kendra's having lied emerges as a second accusation connected to the first one and Kendra's strategy of defending herself by first claiming that the (possible) lie – the asser-

tion that Melanie didn't have a phone – was told long ago, a kind of extenuating move, and then declaring it to be joke:[12]

> KITTY: I thought Melanie didn't have a pho=ne. ... But I'm supposed to call Melanie's mo=m.
> KENDRA: [(SIGHS)] I didn't want you to have her phone number.
> KITTY: Yeah, another lie. Remember?
> KENDRA: That was a long time ago.
> KITTY: Now I'm supposed to re- -- And I'm supposed to trust you every .. [but how am I supposed to know when] you're telling the [truth].
> KENDRA: [That was a joke. That was a joke.]
> KITTY: and I'm supposed to read your mind and know what you're joking about.
> KENDRA: .. No= but, it was a joke.

This passage is closed by Kitty's generalising statement:

> You know Kendra, -- I don't know how many times I gotta tell you, once you lie=, ... once you lie, you lose th- .. complete trust.

Generalising statements of this type are frequently used as a means of closing a topic.

After a short interlude about an unconnected topic which Kendra had introduced, Kitty returns to the conflict by imposing a punishment on Kendra, which motivates Kendra to repeat her denial:

> KITTY: .. You're not going anywhere for a while.
> KENDRA: Go=d, I didn't spend the night, that's what makes me so m=ad, I'm grounded for nothi=ng. ... I was home all night last night, I came home and told you. […]
> Grounded for a month, < yells> and I didn't do nothing.

These repetitions of her denial are a characteristic element of her defence strategy, as the following example shows:

> KENDRA: < addressing Curt > Let me tell you right now, I did no=t spend the night. ... .. This is stupid. […] This isn't fair, I'm being punished for no reason. Fucking stupi=d.

The sequencing pattern she follows can be described as "insisting" (on her innocence).[13] The repetitious element of "insisting" could be considered a doubtful strategy, but, in any case, it demonstrates the seriousness of the denial and the tenaciousness of the defendant.

---

[12] Later on in the exchange her mother even declares this to be the primary problem: "You been caught in a lie Kendra, and that's the whole thing. That's why you're pissed."

[13] For an analysis of the practice of insisting, see Hundsnurscher (1981).

Another interesting type of move which Kendra uses on several occasions is the conditional I mentioned earlier on:

> ... I think if I spent the night I would've took a shower.
>
> See if I spent the night Mom, you'd think I'd still be in my uniform?

With these moves she tries to introduce additional evidence for her having stayed at home. They do, however, not impress her mother, who disqualifies these moves as her trying to "cover her tracks".

What my discussion has been trying to demonstrate is how the sequencing pattern "accusation & answering the accusation" structures the whole exchange and thereby contributes to its coherence. The main features of Kendra's answering strategy consist in her repeated denials and her various (fruitless) attempts to provide evidence for her having stayed at home. It is remarkable that she does not insist on her (possibly) strongest source of evidence, namely phoning her friend's mother. Kitty's counter-strategy employs the additional accusation of lying, which is difficult for Kendra to counter. A relevant aspect of this exchange is that it is obviously played out against a longer history of conflict, which is indicated by Kitty's summary: "Me and Kendra got in another fight". This store of common ground knowledge (and emotion) obviously plays a major role for the moves that characterise this exchange and, more generally, for parts of the organisation of family life.[14]

13.4 Making plans together

Another type of dialogue that is characterised by certain speech act sequences are planning dialogues ("deliberative dialogue"). Typical everyday planning situations include planning a vacation, planning an outing, planning a birthday party etc. But there are also more serious planning occasions in politics, economy, and academia, e.g., planning a new political programme, planning the digitalisation of a firm, or planning the reform of a curriculum. The latter types require a wider array of topics, a larger number of participants, and maybe a longer period of planning, but many of the basic moves are similar to those in the more mundane everyday ones.

---

[14] An interesting aspect I did not discuss is the involvement of other members of the family, e.g., Kendra's stepfather, who obviously tries to stay aloof, and her sisters Sabrina and Marlena. When Marlena takes sides with Kendra, her mother explicitly tells her to back off: "<Kitty yells> Hey, stay out of it. ... Stay out of it, you're not involved in this."

For many of our everyday activities we have well-practiced routines or prescribed methods on which to rely in our choice of individual acts and strategies. However, in some cases we have a practical problem: we do not know what to do. In such situations we may ask someone for advice or we may decide to deliberate together what to do and to develop a plan of action together. The latter type of communication I shall call a *planning dialogue*.

Planning dialogues form a family of forms of communication which are characterised by a number of typical speech acts, speech act sequences, and strategies.[15] One of the basic speech act sequence types that accounts for much of the coherence of such conversations is the sequence: "making a suggestion & discussing the suggestion" with its various extensions.

The starting point for dialogues of this type frequently consists in the common knowledge of a practical problem or the explicit formulation of the practical problem to be solved. Some of the practical problems can have the following form:

(i) A and B are in a disagreeable situation and do not know how to solve this problem.
(ii) A und B want to reach a certain aim but do not know how to accomplish it.
(iii) A und B see various paths to attain their aim but do not know which path to choose.
(iv) A und B have decided to choose a certain course of action but do not know how to implement the relevant actions.

Depending on the kind of practical problem to be solved, we can distinguish different the planning situations and also different phases within a planning dialogue. For instance, a detailed planning dialogue may start with a problem of type (i) or (ii), then move to a situation of type (iii), and finally end by discussing a problem of type (iv). Common knowledge of the practical problem at hand may be given, but in some cases the precise nature of the problem will first have to be identified and communicated to the participants. Therefore, early phases of a planning dialogue will sometimes be concerned with the analysis and clarification of a practical problem. And in some cases, a clarification or even change of the projected aim may only emerge in the course of the planning process itself.

As I already mentioned, a fundamental move in planning dialogues is the suggestion of a course of action. The point of "suggestions" consists mainly in presenting possible actions for evaluation and discussion. This role in planning dialogues can be represented as a number of commitments under-

[15] For a more detailed analysis of this type of dialogue, see Fritz (1982, ch. 8).

taken by uttering a suggestion, as follows:

> A can suggest a common course of action x by uttering "Let's x".
> In doing so, A undertakes the following commitments:
>
> (i) A assumes that x-ing is an option to solve the practical problem,
> (ii) A assumes that x-ing is feasible,
> (iii) A assumes that x-ing is acceptable (morally, psychologically),
> (iv) A evaluates the suggested option positively,
> (v) A is prepared to participate in in the proposed activity.

Depending on the relevant commitments, suggestions come in different *strengths*, e.g., tentative suggestions (i), confident suggestions (i) to (iii), and enthusiastic suggestions (i) to (v). These types of suggestion are sometimes indicated by the choice of expression or the type of intonation. Tentative suggestions are (sometimes) formulated by uttering *Maybe we could x* or *Couldn't we just x?* and in addition hedged by saying *But that's just an idea* or something of the kind. *What about x-ing?* could signal a medium grade of confidence, whereas *I'd love to x* commits to the full set of commitments.[16]

In a way similar to the commitments of accusations discussed in the last section these commitments are related to an array of relevant consecutive moves in the dialogue. In the case of suggestions, it is mainly the *objections* voiced against and the evaluations pronounced concerning suggestions that are characteristic of relevant responses, apart, of course, from outright acceptance and rejection of the suggestion. Take the following examples:

(12) You must be joking. (i)
(13) I don't think that is an option. (i)
(14) That sounds nice, but it is totally impracticable. (ii)
(15) We can't do that. It's illegal. (iii)
(16) That's a brilliant/ridiculous idea. (iv)
(17) Are you really prepared to take part yourself? (v)

In many cases the objections and evaluations will be supported by arguments, e.g., by giving reasons why the proposed course of action cannot be realised, why it is morally unacceptable, or why it is stupid or dangerous (or something of the kind). The types of objections and the dimensions of evaluation considered relevant depend on the kind of action proposed and the kind of projected aim of the action(s). So we can have a vast field of different objections and evaluations, the relevance of which will sometimes have to be clarified in the planning dialogue itself.

---

[16] Other formulations of suggestions include: *Why don't we …*, *How about …* and performatives like *I propose we …* and *I should like to propose/suggest …*

A frequent type of third move in this kind of dialogue is the defence of a given suggestion against (different) objections and negative evaluations. This can lead to several rounds of objections to and defence moves for the *same* suggestion. Other sequences of these types of move include the comparison of different options and the discussion of several proposals according to similar criteria. This can lead to several rounds of *different* suggestions and discussions of similar structure (suggestion$_1$ + objection to suggestion$_1$ + defence of suggestion$_1$ + rejection of suggestion$_1$, suggestion$_2$ + objection to suggestion$_2$ + defence of suggestion$_2$ + rejection of suggestion$_2$ etc.)

Additional moves in this type of dialogue include conditional suggestions (*If we have enough time, we could …*), asking for and giving further information on a suggestion or objection, explaning a suggestion or an objection, and elaborating on the consequences of a proposal. An interesting further extension of these types of moves consists in narratives showing that a proposed course of action was successfully followed in the past or why a proposed course of action is not practicable, dangerous or subject to other problems. Thus, we see how the basic pattern "suggestion & evaluation/ objection" can be extended to produce longer planning dialogues in many variants.

The use of this basic repertoire of moves, which accounts for the coherence of many exchanges of the planning type, sometimes also raises matters of strategy. A typical kind of response to a suggestion made in academic committees consists in making a not too enthusiastic positive evaluation ("a very interesting idea") and following this up by a (much more) grave objection.

This strategy is (sometimes) meant to mitigate the objection. Another interesting phenomenon is the strategy of making objections to all suggestions presented without making a suggestion of one's own. This is sometimes criticised as being uncooperative, but, of course, there may be very good reasons for objecting to all suggestions considered bad. An interesting variant of this type of sequence is the technique of "brainstorming", where ideas for possible courses of action are accumulated without discussion. This is supposed to produce particularly creative proposals.

As an example of a simple everyday version of this dialogue type I give the following face-to-face conversation among four roommates (students, women ages 20-21), recorded in a shared apartment:[17]

> <After a period of small talk, mainly between Dana and Kelly, who both state that they are tired, Nancy opens up a new topic.>
>
> NANCY: So what's the pla=n.
> DANA: .. I don't know. I just was -- Went shopping with .. Ed and ... his roommate, they said they'd come over here.
> NANCY: ... Mm.
> DANA: You know. ... If we wanna just hang out and drink. ... Whatever.
> NANCY: Okay. That [sounds .. cool].
> DANA: [I think it'd be cool], cause I don't feel like going anywhere.
> NANCY: [Yeah]. That sounds good.
> ARIANNA: [I know].
>
> <a few minutes later>
>
> DANA: So I don't know what they're doing. But I can call, and find out. If they wanna come over.
> NANCY: .. Okay.
> DANA: In a little while. I wouldn't mind, I just wanna hang.
> NANCY: ... Yeah.
> DANA: [Ha=ng and talk].
> NANCY: [We don't have ca]ble.That's the only thing that sucks.
> DANA: (makes a click sound)
> NANCY: [So we don't have anything] --
> DANA: [We have a ster]eo.
> NANCY: Yeah.
> DANA: [You know],
> ARIANNA: [That's true].
> NANCY: [You're right]. That's [like],
> ARIANNA: [That's] fine.

This sequence is opened by Nancy's question signalling that she does not know what – if anything – is planned for the evening and that she would like

[17] Like the transcription in the last section, this transcription is taken (slightly simplified) from the Santa Barbara Corpus of Spoken American English (SBC050). https://www.linguistics.ucsb.edu/sites/secure.lsit.ucsb.edu.ling.d7/files/sitefiles/research/SBC/SBC050.trn.

to know. Dana's response consists of two parts. With the first she tells Nancy that she has no plan, and the second is a minimal report of earlier activities, which signals a possible development of the evening: Ed and his roommate "coming over here". Nancy responds by a kind of "keep going signal" and Dana continues by making her suggestion more explicit, however in a rather tentative fashion, by using an (incomplete) conditional form ("if we wanna just …") and adding, somewhat noncommittally, "whatever". Nancy agrees with the suggestion by giving a positive evaluation ("Okay. That sounds .. cool"), talking in parallel with Dana, who also evaluates the suggestion herself ("I think it'd be cool") and adds as an explanation that she does not feel like going anywhere. This explanation could also have the additional function of justifying her – probably not very exciting – idea. Nancy repeats her positive evaluation, while Arianna, in parallel, reacts to Dana's explanation by stating that she knows Dana's reason for "just hanging out". This could be considered to be a less than enthusiastic reaction.

The second round of this conversation is linked to the first by the idea mentioned earlier on that the boys could "come over". Dana agrees to call them – with a small delay – and repeats her original suggestion in a slightly modified version by stating her personal preference that *she* just wants to "hang", adding "hang and talk". At this point Nancy makes an objection by stating that they don't have "cable" (i.e., a cable radio), so they can't hear music, and explicitly deplores this ("that sucks"). Dana responds, dispelling the objection by stating that they can use their stereo (radio) instead. This addition to her original suggestion is then also accepted by both Nancy and Arianna, which concludes this round of the exchange.[18]

What is remarkable about this exchange is the form in which the main suggestion is made – first in the conditional form and later as a statement of personal preference – and the way in which a secondary suggestion – the boys "coming over" – is linked to the first.

As I mentioned at the beginning of this section, the basic elements of this dialogue type are also used on much more formal occasions, e.g., in the case of public participation in the introduction of administrative measures. The following example is taken from a report on the standard procedure of introducing an item of traffic regulation in an English town.[19]

---

[18] This detail of a dated technology indicates the period when the recording was produced.

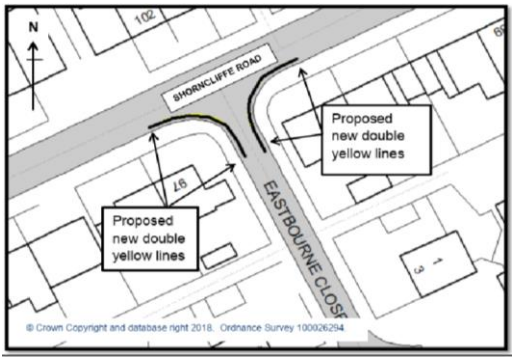
[19] Coventry City Council. Public report. Cabinet Member Report. 12th August 2019. (https://edemocracy.coventry.gov.uk/documents/s44788/Objections to Proposed Waiting Restrictions Variation 8.pdf). (24.10.2021).

(18) <Problem>

Residents' concerns raised regarding access issues, especially for waste services and Ring & Ride services due to parking at junction (Eastbourne Close/Shorncliffe Road).

<Proposal>

Double yellow lines (no waiting at any time) proposed for junction protection.



<Objections>

The proposed yellow lines would be outside our house. With [describes personal circumstances] we would find this a real inconvenience. One of the reasons we bought the house [no. of years] was due to the good parking outside the house. A drive would be very hard to do [reasons]. We have never seen an accident or near miss so cannot think this is the reason for the double yellow lines. I do not oppose the yellow lines on Eastbourne Close, I just feel they do not need to be at the front of our house. We are happy for it to be at the side, on Eastbourne Close.

<Response to objections>

In response to the objections the proposed length of double yellow lines has been reviewed. It is considered possible to reduce the proposed length of the double yellow lines on Shorncliffe Road by approx. 1m on one side of the junction and 1.5m on the other. This would result in a space of approximately 5 metres (car length) remaining outside both corner houses, starting at the boundary with their neighbour.

The coherence of this use of the interactive procedure is based on several factors:

(i) The general dialogue pattern "proposal & objection to proposal & response to objection",
(ii) mutual knowledge about the situation at the junction,
(iii) continuity of the topic (traffic problems at the junction),

(iv) coreference relations between expressions in the different sections of the exchange (e.g., *the proposed yellow lines*, street names),
(v) assumptions concerning what counts as a valid objection,
(vi) mutual knowledge about the roles of the participants in this procedure.

Concerning these factors a few remarks are in order. The utterances intended as objections receive their function on the basis of their sequential position after a proposal and based on common ground as to what is a possible valid objection. Taken in isolation from this sequential position the sentences in section <objections> could simply function as statements, some of them as counterfactual statements (e.g., the statement of reasons why they bought the house, *We would find this a real inconvenience*). And if the person(s) objecting to the proposal had mentioned the fact that their house was painted red, this would have been difficult to understand and to accept as an objection. As for the mutual knowledge about the situation at the junction, the little map contributes to clarifying the situation (e.g., the intended position of the yellow lines), thereby securing mutual knowledge. To understand the general relevance of the proposal, the objections and the response, one has to know the rights and responsibilities of the participants, e.g., the right and responsibility of the city council to implement traffic regulations.

This example demonstrates again a point I mentioned at the beginning of this chapter: sequencing patterns play a fundamental role in the creation and monitoring of coherence in dialogue, regularly interacting with other factors like common ground and topic continuity. The latter aspect of coherence in dialogue will be dealt with in the following chapter.

A final point to take into consideration here is the fact that there are dialogue types or dialogue genres that provide typical "programmes" of sequencing patterns and (types of) topics, the same as we observed in the case of text types/text genres. Dialogue types of this kind are, for instance, oral controversies, deliberative dialogues, negotiations, the exchange of narratives, gossiping, telling jokes, interviews, teaching discourse, examinations, shopping discourse and many others.[20] Knowing the basic types of moves and sequencing patterns of these dialogue types is a prerequisite for making coherent and appropriate contributions to dialogues of the respective types.

[20] For a detailed analysis of gossip, see Bergmann (1993).

14. Topic management in dialogue

14.1 Topics in dialogue

Together with the practice of turn taking, certain forms of topic management belong to the most characteristic elements of dialogues as compared to monological texts. From the point of view of a dynamic text theory these forms of topic handling are particularly interesting as the stepwise collaborative production of topics is a fundamentally dynamic procedure. A basic question to explore is the following: How is it possible to create and monitor the topical dynamics of a conversation while at the same time preserving and recognising topical coherence?

Probably the most spectacular phenomenon in this connection is when, with two or more persons interacting in the production of the flow of conversation, topic developments and topic structures *emerge* that were neither planned nor intended by the individual participants. To describe some of the kinds of moves that may lead to such an outcome in the course of quite natural, coherent topic talk will be one of the aims of this chapter.

It is, however, worth noting that not all types of dialogue are favourable to such a kind of topic development and that, on the other hand, there are also forms of written communication, e.g., written controversies or dialogues in social media, that produce quite unforeseeable topic developments. Many examinations, for instance, belong to a type of dialogue where a certain set of topics is more or less fixed in advance and where these topics are sequentially processed in the course of the colloquy. In such a context it might occur that when the examinee "strays" from the narrowly conceived topic the examiner will guide them back to the intended topic path. Therefore, in this kind of dialogue a form of topic *continuity* can be considered the default mode. But even everyday conversations are frequently fairly standardised and focused on a small range of topical issues, the movement from one subtopic to the next in a cluster of topics being highly predictable, which makes for an appearance of continuity. If, for example, in the context of an important event in a small community, e.g., the election of the mayor, members of this community meet, it is likely that they will talk about the candidates, their qualification, their history etc., making use of the common knowledge available in the community to produce a smooth flow of topics.

From the point of view of the coherence of dialogue it is useful, as I mentioned before, to treat the topic as an aspect of dialogue that is *managed* and *achieved* sequentially by the moves of the interactants against a background

of topic knowledge that the participants bring to this task. This approach, which I also used in my treatment of text topics in the first part of this book, has the advantage of combining the general focus on the *activities* of the participants with the focus on the topic structures emerging as part of these activities. The usefulness of this approach has been emphasised by authors from quite different research traditions, e.g., Bublitz (1988, ch. 2), Schegloff (1999, 53), Schröder (2003, 83), and others.

14.2 Introducing a topic

Where does a topic begin and how does it begin? In many cases, for instance, when speakers explicitly indicate the beginning of a topic, this is quite obvious. However, sometimes topics in conversation seem to begin quite inconspicuously so that their beginning is only noticed in retrospect. This can be the case when there is a stepwise shift of topic reaching a point where the current topic seems to merge into what turns out to be a major new topic. So the new topic seems to "just happen along" (Sacks 1992, Vol II, 352). Some of the processes leading to such a development will be discussed in the next section. In the present section I shall now consider some of the devices used to introduce topics explicitly.

In some cases, especially in formal contexts, for instance in lectures, the speaker will directly and explicitly *announce* the topic with expressions like the following:

(1) I will start with some general remarks on graph theory.

Somewhat more informal are utterances like (2)

(2) I have to tell you about the new book I am reading.

In cases like these the act performed ("announcing") serves exclusively to introduce a topic, whereas in many other cases various types of acts have a double function, one of which consists in contributing to topic management.

Maybe the most frequent method of introducing a new topic consists in using a question to propose or proffer a topic.[1] This kind of pattern is normally used early on in an exchange or at a point where another sequence has been closed. An example of the first is:

[1] "proffering" is the term Schegloff uses in his analysis of this type of pattern (Schegloff 2007, ch. 8). On the use of questions for the introduction of new topics, see also Riou (2017).

(3) A: Hello
B: Hello
A: How was your weekend?

At this point, B has the option of either taking up the proposed topic by, for instance, mentioning an interesting activity they took part in or of indicating that they don't want to talk about the weekend by saying something like "Oh, it was alright". In the latter case, the topic does not take off and A could either let it go or make another try to get it going. In the first case, B has the opportunity to tell a story about exciting activities and A can join in by, for their part, telling something about *their* weekend. Thus, the topic of weekend activities is established.

A second type is exemplified by the following:

(4) A: <finishs telling a story about his elder brother.>
B: Ah yes, very nice.
A: Yes.
<1 second pause>
B: And what is your *younger* brother doing?

After A has finished telling his story a transition relevant position for topic organisation is reached. B makes a typical recipient remark and A acknowledges this. At this point a short silence intervenes.[2] By asking the subsequent question, B cooperatively contributes to the continuity of the conversation, providing the opportunity for A to give some information about his younger brother's work, his family life or his travels etc., which launches a new, albeit related, topic. Thus, introducing this related topic achieves a kind of (minor) topic shift.

A variant of this type is the case where the narrator himself uses the closing of his story to offer one of his recipients the opportunity to take the floor and present a topic of their own. The following example is taken from a conversation between three academics:[3]

(5) A: <finishs telling the story of how he exposed a fraud concerning a valuable object being offered to the British Museum.>
Anyway. That was –
B: So you exposed it
A: I exposed it. Yes. But they can't get it BACK. – Well I mustn't go on boringly talking about ME, what are YOU doing Geoffrey?
C: um uh much the same old things.

---

[2] On the role of silences for topic development, see Maynard (1980, 265-268).

[3] Svartvik/Quirk (1980, 237) (transcription slightly simplified); see also Bublitz (1988, 69).

In this particular case C ("Geoffrey") seems to be reluctant to take up the proffered topic and talk about his personal situation, which he indicates by using hesitation markers, and, in the following exchange, it takes A several attempts to bring C to tell them some details about his teaching job.

Another type of topic introduction uses the common knowledge and the joint attention of the interactants at a given point in a conversation. A typical kind of example is the following, which could occur during a meal with friends:

(6) A: I rather like this wine.
B: Yes, it's a Barbera, from the Piedmont. They produce some very nice wines there. <B continues with some remarks on wines from this region.>
A: <A tells a story about drinking wine in Puglia.>

The guest A compliments his host on his choice of wine, which is in the focus of the attention of the participants. By doing so, he not only, at one level, makes a polite move but he also opens up a possible topic, namely the topic of wine. B takes up the topic and elaborates on it. Now A himself has the opportunity to continue this topic by contributing some relevant remarks of his own or a story about wine.

Basically, all things, events etc. that are or can become objects of joint attention in a given situation can be remarked on and thereby proffered as a topic.[4] And what is worth noting once more about this practice is that mentioning a potential topic can be an aspect of many different speech acts, e.g. questions, statements, compliments, requests etc. What makes such a case of mentioning an introduction of a topic is its being taken up and continued for at least a short period.

Whereas in the last few examples A provides an opportunity for B to talk about a topic, there are also other cases, where A creates such an opportunity for themselves. Typical moves of this type, which I mentioned earlier on, serve to prepare the telling of a story, like in the following examples:

(7) A: Guess what happened to me today. …

(8) A: Something weird happened to me today. …

Besides requesting attention for the telling of a story, these formulaic expressions also serve to prepare a *kind* of topic by announcing that a remarkable event will be told. In both cases, the speaker will, after this preparation, nor-

[4] On the technique of using the "occasionality", i.e., the specific, known-in-common context of a conversation, as a resource for generating topics, see Adato (1980); see also Bergmann's concept of "local sensitivity" (Bergmann 1990).

mally continue straightaway with their telling of the story, thereby elaborating the topic.

Depending on the position in the course of the conversation, these devices can be used to open up a first topic in a conversation or to change the topic, sometimes quite radically. In the latter case, the use of such devices indicates that the speaker is aware of the topic change and tries to "normalise" it, thereby contributing to maintaining coherence.

14.3 Shifting the topic

As I mentioned at the beginning of this chapter, in many cases topic changes occur incrementally, step by step. Such occurrences have been termed topic shading, topic shift, topic drift or stepwise transition.[5] A point worth noting is that such topic shifts are frequently related to particular sequencing patterns like telling a story, giving an example or giving an explanation. In the following passages I shall give a number of examples of moves where topic shifts tend to occur.

Generally speaking, *stories* frequently open up opportunities for topic shifts. Take, for instance a general conversation about national stereotypes. In the course of talking about this topic, person A states that there is a stereotype that Scottish people are generally stingy. In order to prove the unfoundedness of this stereotype, person B tells a story about meeting many quite generous people when living in Scotland. In the course of this story B may mention various properties and actions of these people that have nothing to do with stereotypes in general and stinginess in particular. And A may ask B questions about B's stay in Scotland that are also not relevant to the topics of stereotypes and Scottish meanness in particular. So what we find in this kind of conversation is a process of continuous topic shift from a rather general topic to a specific aspect of this topic and from there, via a narrative, a stepwise transition to topics that are related from step to step but no longer directly related to the original starting point.[6]

The following example also shows a shift from a general topic to a personal one. Two girls, Jennie and Lisa, are talking about the quarter system in

---

[5] See Schegloff/Sacks (1973, 305), Bublitz (1988, 124ff.), Hobbs (1990), Jefferson (1984).

[6] Similar topic developments can be observed in threads of postings in social media, e.g., on Twitter. An example from the Chaucer mailing list is discussed in Bader (2018, section 6.1).

college in the USA – as opposed to the semester system – and how hard it is to keep up with the work:[7]

(9) 1 Jenny: It's really pressur [ing]
2 Lisa: [They] move so fast
3 Jenny: um humm
5 Lisa: You jus' – you know GAWD [you jus'] res' fer a DAY an'
6 Jenny: [Oh yeah]
7 Lisa: you're way behind [it seems like]
8 Jenny: [Oh ye::ah] An I been PLAY:in aroun'
9 too much
10 Lisa: Ri:lly?
11 Jenny: Cuz I'm USE ta PLAY:in' …

This extract of a transcription begins with Jenny's general remark on the pressure exerted by the quarter system and Lisa continues by specifying in general terms what the pressure consists in ("They move so fast"). Jenny agrees to Lisa's description by making backchannel utterances ("Um humm" and "Oh yeah") and then, in utterance 8, she continues by talking about a *personal* problem with the system ("An I been playing around too much"). Like "resting for a day" (line 5) "playing around" is an example of the kind of behaviour that is detrimental in the quarter system. Thus, she is still talking about problems within the quarter system but has shifted to the personal level and a problem of her own. This kind of movement from general topic to individual (and sometimes back again), especially in the course of *giving examples*, is quite a frequent pattern of topic shift that is well within the boundaries of coherent topic talk.

Another pattern of topic shift is connected to the *giving of explanations*. The following is an example of this kind of procedure:[8]

(10) <Two tenth-grade friends, Robert and Tim, are talking about girls, the upcoming ring dance, an important event for them, and about going out tonight.>

Robert: … but then I really don't feel like going tonight, and I'm thinking Mom's not paying me for ring dance and so I have to save my money. So I –
Tim: How much do you have?

< For a few turns they talk about money. And then they continue talking about girls and drinking at parties.>

[7] Maynard (1980, 271f.). In this transcription capital letters indicate emphasis and ":" indicates lengthening of vowels.

[8] Hobbs (1990, 17); the full transcription is given in Dorval (1990, 326-340).

In this stretch of talk Robert's explanation of why he does not want to go out introduces the topic of money and this is taken up by Tim's question, and in this way the topic shift to money is performed quite naturally, creating a new focus of the conversation for a number of turns. As Hobbs (1990,11) noted, "in casual conversation, the explanation sometimes turns out to be more interesting than the explained, and the conversation somehow never gets back to the original topic". In this particular case, where girls and parties are the general topic of this conversation all along, the topic shift is only temporary, so what we have here is a kind of *digression,* another frequent type of topic management in everyday conversations.[9]

In fact, digressions are such a remarkable feature of topic development that we find certain expressions ("displacement markers") routinely used to explicitly indicate the opening of a digression (e.g., *by the way*) or the return to the original topic (e.g., *anyway*, *coming back to*, *but back to*, *going back to*).

Whereas in many cases these kinds of topic shading "just happen" due to the various and dynamically shifting interests of the participants, there are also cases where these techniques are used intentionally to either lead up to a certain topic or to steer away from a disagreeable or otherwise difficult topic. Leading up to a topic by several steps may appear a good strategy in cases where a more direct approach may seem embarrassing or otherwise awkward. By preparing the presentation of a topic in this way the speaker may veil their interest and the degree of importance this topic has for them. Sacks described the technique as follows: "Now, the character of the stepwise movement for topics is that if you have some topic which you can see is not connected to what is now being talked about, then you can find something that is connected to both, and use that first" (Sacks 1992, vol. II, 300). Steering away from a difficult topic stepwise is an alternative to closing down the topic abruptly, which could appear brusque, unfriendly or impolite. One version of this technique consists in picking up secondary aspects of the problematic topic – what Jefferson calls "ancillary matters" (Jefferson 1984, 202) – and continuing talk on these, thereby slowly moving away from the original topic. What this technique accomplishes – among other things – is to preserve (the appearance of) topical coherence while traversing remarkable topical distances.

The examples presented in this chapter show that, on the one hand, there are pre-existing and ready-made clusters of topics and subtopics, which can

[9] On the structure of digressions in conversations, see Dascal (2003, 213-243), Bublitz (1988, 94-124).

be used for stepwise topic movement by following their lead, but, on the other, that frequently topical connections are created ad hoc by using certain speech acts with a secondary function as devices for proposing a – more or less – related topic. If these devices are repeatedly used, the topic may go off on a tangent, leaving the participants with the question how they got there.

14.4 Closing a topic

Some topics in conversation are closed by external interruptions like a telephone call, the beginning of a lecture or the fact that one of the participants has to leave. And there are cases where a participant explicitly suggests or abruptly demands that a topic should be abandoned ("Let's drop this topic."). More polite versions include the use of expressions like "I don't want to keep you any longer ...". Other closings are a side-effect of unintentional topic shifts of the kind mentioned in the last section. Again other closings seem to come about by the topic's simply petering out when the parties to the conversation have nothing more to say about the matter. But in most cases some kind of collaborative "work" has to be performed by the interactants to successfully close down a topic.

Of the many structural points of a conversation at which closing a topic occurs I shall mention two here, the closing of a whole conversation and the closing of a longer sequence like, for instance, a narrative.

The following type of topic closing is intimately connected to the closing of the whole conversation. If, for instance, the point of the conversation is to plan a shared activity, the topic of the future activity can be closed by making final arrangements for the meeting etc.:

(11) A: OK, so we'll meet at seven, right.
B: OK

This brief exchange not only serves to close the topic but simultaneously prepares the closing of the whole conversation. Of course, another topic could be raised at this point, for instance by B's saying: "What I meant to ask you …" or something of the kind, but basically moves like (11) open up closings.[10]

Not necessarily the end of a conversation but at least the closing of a topically coherent stretch of talk can be brought about by the closing of a story.

[10] Various techniques of "opening up closings" of conversations are discussed in Schegloff and Sacks' famous paper of this title (Schegloff/Sacks 1973).

Stories frequently serve to make a point about a given topic, e.g., a stereotype of the kind mentioned earlier on. In such a case the story's ending can be indicated by the narrator's mentioning this topic, e.g., by drawing the moral of the story ("So you see, the Scottish are not stingy at all"). This move, while concluding the story by the claim that the speaker has made their point, can also indicate that the topic can be closed. As in other cases, this option has to be acknowledged by the interlocutor, but it may also not be accepted and the other person may start on a topically related story which either confirms this claim or throws doubt on it. Generally speaking, various types of summarising moves can function as indicators of topic closing. The following is an example of this type:[11]

(12) <P and M are discussing P's unemployment>

P: <breathes audibly> But I think it'll iron itself out.
M: I sure hope [so
P: [I'll see you Tuesday

In this excerpt P indicates a possible end of topic by giving a summary statement of her attitude, M agrees and expresses her empathy, and P segues into a typical closing move. So, again, the ending of the topic is a collaborative effort.

The closing of a topic, in short, can come about in many ways, frequently by the use of quite subtle procedures facilitating the transition while retaining coherence and the flow of talk.

14.5 Concluding remarks on topic talk

In this chapter, I emphasised aspects of topic management that are characteristic of dialogue. The following aspects emerged as particularly noteworthy:

(i) Topic management in dialogue is a collaborative achievement, with turn-taking as a basic factor.
(ii) In addition to common knowledge of topic structures, co-presence and shared attention provide "occasionality" as an important resource for the production and management of topics in dialogue.
(iii) For routine topics, topic continuity is the default mode.
(iv) Interactants apply various subtle topic shifting procedures.
(v) Topic structures frequently emerge unplanned.

[11] Excerpt from Holt (1991, 131).

(vi) Functional sequencing and topic management interact in the production of coherent dialogue.

At the same time, dialogue and monological text have a number of fundamental aspects of the use of topics in common. In the first place, known-in-common topic structures and topic-related common knowledge are basic resources of dialogue, very much like in written communication. Furthermore, the interaction of functional sequencing and topic management plays an important role both in dialogue and in monological text, albeit with certain differences in sequencing patterns. And finally, as for procedures of topic change, routine patterns are also used in writing, however without the methods provided by turn-taking. Thus, we come to the conclusion that topics and their interaction with functional sequencing are basic factors of coherence in discourse, which are, however, *used* in partly different ways, adapted to the specific tasks of dialogue and monological text.

15. Conclusion

In their social lives, speaker/writers try to master communicative tasks of different size and complexity, from referring to an object to convincing someone of the truth of a claim and to telling them the story of one's life. For their performance of such tasks in the flow of discourse, speaker/writers have available a cluster of communicative resources, e.g., established sequencing patterns for linguistic acts, including their language-specific utterance forms like connectives and anaphoric devices, their knowledge of current topics, and varying degrees of common ground. In performing various series of communicative acts by using these resources, they follow a number of communication principles, including the principle of recipient design, the consistency principle, the principle of coherence and various politeness principles. Hearer/readers use the same resources and, in addition, their faculty of drawing inferences to make sense of the discourses in which they participate.

Among the principles mentioned just now, the coherence principle has a special status in that it is closely related to the comprehensibility of contributions to discourse, as a fundamental aspect of discourse comprehension consists in "seeing the connections" of contributions to a discourse. This principle and the methods of its implementation have a number of interesting but also disconcerting properties, which I shall now briefly summarise.

The coherence of a discourse is an achievement of the participants to the discourse. In a fairly obvious way, this is true of conversations, where creating coherence is a collaborative act. But it is also true of monological texts, however, in a different sense. Speaker/writers usually attempt to construct a coherent discourse, thereby adjusting their contribution(s) to the (possible) recipients, making use, for instance, of sufficient common ground, enabling the recipients to "see" the coherence. On the side of the hearer/reader, it is the *presumption* of coherence, a counterpart to the principle of coherence, that can be seen as both a resource of and a requirement for comprehension and interpretation.

The principle of coherence is both a *normative* principle and a principle that is *routinely* followed. This somewhat ambiguous status explains why, on the one hand, deploring the incoherence of a text or any contribution to a discourse may be considered a relevant complaint, whereas, on the other, most of the time the coherence of discourse is simply assumed, and minor discrepancies are routinely "repaired".

This *elasticity* of the demands on coherence is due to the vagueness of the rules and the flexibility of the methods applied in creating connected discourse. This is one of the reasons why it is so difficult to formulate strict

constraints on coherence, apart from some fairly general ones I discussed in chapter 8. So, in a way, it seems more enlightening to ask for the resources drawn on in constructing coherent discourse and the methods of their application than asking for necessary conditions of coherence. And, of course, there are many forms of "anarchistic" and disruptive text production, e.g., in literary discourse, where violating principles of coherence is a specific resource for literary effects.

One of the most interesting issues in the exploration of coherence is the *interaction* of the resources used in creating and understanding coherent discourse. I shall now briefly recapitulate some of the results of my discussion of individual resources and then go on to describe some of the interactions.

It is obvious that *common ground* is the background to all kinds of communicative acts, from reference to objects to the explanation of actions and the production of a coherent narrative. If we define the *context* of an utterance as the common ground "active" at the point of this utterance, the most remarkable property of this context is that it both contributes to the understanding of the utterance and is *changed* by the utterance. Therefore, *managing* the dynamics of common ground (as a speaker/writer) and *monitoring* it (as a hearer/reader) is a fundamental precondition of successful coherence management.

From the point of view of a dynamic action-theoretic approach to discourse coherence, the established *act-sequencing patterns* routinely used by the participants take pride of place in the cluster of communicative resources. Basic patterns are simple two-part sequence types like "making a claim & supporting the claim", "making a request & justifying the request", "question & answer", and "criticism & rebuttal". These and many other basic patterns can be combined in various forms to incrementally build up complex act structures, some of which are hierarchically structured, providing the potential for complex communicative activities. In the words of Barbara A. Fox: "presumably the structure one finds of a finished text or conversation is a by-product in some way of the many smaller, locally situated units produced in real time" (Fox 1987, 154). From the point of view of the hearer/reader something similar can be said: "clearly the reading process, which the writer must anticipate with the use of things like anaphoric devices, is locally situated […] and very much a process rather than a product […]" (Fox 1987, 155). Slightly more comprehensive patterns are represented by what I call "minimal text types", e.g., paragraph-sized simple descriptions, minimal narratives, and simple argumentations. These basic text types can also be combined and can be extended by inserting various types of other patterns. In the case of narratives, these patterns may include explanations, justifications, stretches of dialogue, and evaluations. And finally, there are full-blown text

types like (various types of) news reports, concert reviews, instruction manuals, cooking recipes, medical leaflets, research papers, sermons and many others, which provide guidelines for the use of all relevant resources, i.e., sequencing patterns, topic structures, and knowledge management, thereby also providing criteria of coherence. An interesting complication can be found in discourses where a change of text type or dialogue genre takes place.

Topics are difficult to define as a category but they are nevertheless a very real resource of discourse coherence. The knowledge of topics and their related frames of knowledge is used as "material" for discourse and can also be described as a "sequential structuring mechanism" that provides guidelines for the relevance of contributions to the respective discourse. Topics can be analysed in terms of several structures, i.e., network-like, hierarchical, and sequential structures of topics and subtopics, which are all used for the management of topics in combination with other resources. For some text types, topic structures may even be the primary coherence factor. The collaborative creation of emerging topics is one of the most characteristic aspects of some types of dialogues, e.g., informal conversations.

As for the *interaction* of these resources as a source of coherence, I shall now refer to a few examples I discussed in the course of my explorations. (i) Speech act sequencing patterns interact with common ground, as each act changes the context by adding specific new commitments to the commitment store, which then (normally) become common ground. Therefore, act sequencing and common ground management (and monitoring) go hand in hand. (ii) At the same time, performing an act sequence often requires making reference over and again to the same entities, a requirement that can be met by using patterns of coreference. Of these, the anaphoric use of pronouns and (sometimes) other referring expressions also contributes to common ground management by indicating which entities are treated as being common knowledge at a certain point in the discourse. (iii) Coreference is also one of the means of topic management, as creating a topic sometimes, but not always, requires referring repeatedly to the same entity. (iv) As I indicated before, the interaction of speech act sequencing and topic management plays a fundamental role in many discourse types, which mainly consist in patterns coordinating speech acts and topics, e.g., in the case of theatre reviews with their combination of various act types and a typical set of topics, including (brief descriptions of) the plot and the stage setting, (brief narratives about) the biographies of the actors, and (an evaluation of) their performance. This interaction is also noticeable in the case of functionally "flat" sequences like descriptions and narratives, which are mainly based on a sequential pattern of "continuation" and where the coherence of the discourse

is mainly secured by the topic and its development. Furthermore, the commitments undertaken by the performance of a certain type of speech act may open up the space for a certain type of topic, e.g., the topics of responsibility or relevant norms in the case of an accusation. Finally, certain act types can lead to topic shifts, e.g., explanations, where the explanans may be thematically quite distant from the explanandum. (v) Particularly interesting are interactions between text and image in multimodal configurations, where various aspects of text and image interact, e.g., textual descriptions and (parts of) pictorial representations.

In sum, what we see is that basic communicative resources interact in quite diverse and subtle ways in the performance of communicative tasks and the establishment of coherence and that the explanation of coherence has to refer to these forms of interaction (in many cases).

One of the central points I tried to make in the course of my exploration of coherence was that basic communicative resources are used in both dialogue and monologue for the purpose of creating coherence and that it is worth showing these commonalities, without, however, overlooking the specific features of monologue and dialogue.

Fundamental commonalities are certainly found in the use of common ground and general topic knowledge in both monologue and dialogue. And multimodal elements can be used both in monologue and dialogue. Even various sequencing patterns used in monological texts have dialogical counterparts. This is true, for instance, of minimal text types like description, narrative, and argumentation, but also of sequencing patterns like "assertion & explanation", which can be seen as monological variants of "assertion$_A$ & *why*-question$_B$ & explanation$_A$".

Another point worth noting is that, from the point of view of the theory presented in the present book, monologue is in many respects also dialogical. The following aspects come to mind:

(i) Recipient design is a dialogical principle,
(ii) the concept of commitment is basically a dialogical concept,
(iii) common ground is a dialogical concept,
(iv) many monological texts are explicitly produced in dialogical contexts, e.g., in the case of controversies or contributions to social media.

On the other hand, some "monological" features regularly appear in dialogue, e.g., longer passages without change of speaker, which occur in the case of didactic discourse, narratives or extended argumentations, as I indicated before.

The most striking distinguishing features of dialogue, however, are, of course, turn-taking with its many subtle variants and uses and the collabora-

tive construction of topics mentioned above, which can lead to the unforeseeable emergence of quite new topics, topic structures and topic developments. (This does not mean, however, that producing a text will not also frequently be an emergent construction process, both in individual text production and, particularly, in forms of collaborative text production.) These features require special analysis, especially in respect of the coherence of dialogues, as I showed for the practice of topic shifts. Having noted the commonalities of sequencing patterns mentioned above, we have to acknowledge, equally, that many such patterns are characteristically dialogical, e.g., "question & answer", "accusation & excuse", "request & refusal", and "suggestion & objection". But, certainly, "answers", "excuses", "refusals", and "objections" can also be presented as (written) texts in response to the respective initial acts.

Finally, there are, besides the text types/genres mentioned in my analysis, also diverse *types/genres of dialogue*, like, for instance, teaching dialogues, examinations, interviews, sales talk, negotiations, gossip, debates, and quarrels, which provide orientation for conducting coherent dialogues of the respective types. In addition, there are many hybrid forms, e.g., lectures with question rounds and debates with interspersed monological statements.

To sum up, the resources for the production of coherent discourse discussed in the present book, their interaction, and the practices of using these resources in the flow of discourse have proven useful categories for the analysis of problems of coherence in monological and dialogical discourse.

At the same time, my exploration of coherence has shown that theories and methods possessing a too restricted inventory of categories, e.g., Rhetorical Structure Theory with its limitation to "rhetorical relations" and Centering Theory, (mainly) confined to "entity coherence", cannot account for many facts of discourse coherence, e.g., the dynamics of context change and the incremental development of common ground as well as the frequent interaction of sequencing patterns with topic management. This is also true, in a somewhat different way, of many psychological experiments on text comprehension, which frequently use simple propositional models of discourse with a limited spectrum of coherence factors. On the other hand, as I mentioned in section 4.14, it must be noted that these methods have been used for corpus studies, producing quantitative results, albeit within the limits of their theories. Thus, some of the most interesting aspects of discourse coherence still await large-scale corpus analysis.[1]

---

[1] A number of corpus-based studies analysing global coherence structures within an action- theoretical framework were mentioned in section 4.10.3.

An interesting recent development is the production of text-generating models based on machine learning, which have achieved the potential to produce remarkably human-like text. It is therefore not unlikely that machine learning techniques may also improve the detection of implicit coherence relations. The performance of one of these models (GPT-3) was discussed by me in some detail in chapter 11. From a theoretical point of view, the achievements of such models are quite challenging, as the quality of material they produce on the basis of a training on vast amounts of text, discovering statistically relevant surface patterns, visibly surpasses what text-generation models based on formal discourse theories achieve. This fact is, of course, critical, as one of the motivations for the construction of such formal theories lies exactly in their prospect for successful text generation. This development does, however, not affect discourse theories in general, since part of their motivation consists in modelling the human practice of discourse management, which does, as a rule, not rely on a statistical analysis of vast corpora.

Despite the many insights gained in studies of discourse coherence, there are many problems, especially relating to the interaction of different organising principles for discourse, that are still not well understood and that deserve further analysis and empirical examination. This concerns both many types of discourse and the practices of coherence construction in different languages and cultures. Thus, further exploration of coherence in discourse remains a worthwhile enterprise for the future.

References

Adato, Albert (1980), "Occasionality" as a constituent feature of the known-in-common character of topics. In: Human Studies 3(1), 47–64.

Adams. Douglas (1979): The hitch-hiker's guide to the Galaxy. London: Weidenfeld & Nicolson.

Adorno, Theodor W. (1970): Minima moralia. Reflexionen aus dem beschädigten Leben. Frankfurt am Main: Suhrkamp.

[Aelnoth, "Sanctus Canutus"] Gertz, Martin Clarentius (1908-1912) (ed.): Sanctvs Canvtvs Rex et Martyr. In: Vitae Sanctorum Danorum. København: Gad, 27-168.

Aitchison, Jean (1994): "Say, say it again Sam": The treatment of repetition in linguistics. In: Fischer, Andreas (ed.): Repetition. Tübingen: Narr, 15-34.

Alston, William P. (1964a): Philosophy of language. Englewood Cliffs, N.J.: Prentice-Hall.

Alston, William P. (1964b): Linguistic Acts. In: American Philosophical Quarterly 1/2, 138-146.

Anscombe, G.E.M. (1957): Intention. Oxford: Blackwell.

Asher, Nicholas (1993): Reference to abstract objects in discourse. Dordrecht: Kluwer.

Asher, Nicholas (2004): Discourse topic. In: Theoretical Linguistics 30, 163-201.

Asher, Nicholas (2008): Troubles on the right frontier. In: Benz, Anton/Kühnlein, Peter (eds.): Constraints in discourse. Amsterdam/Philadelphia: John Benjamins, 29-52.

Asher, Nicholas/Gillies, Anthony (2003): Common ground, corrections, and coordination. In: Argumentation 17, 481–512.

Asher, Nicholas/Lascarides, Alex (1998): Bridging. In: Journal of Semantics 15(1), 83-113.

Asher, Nicholas/Lascarides, Alex (2003): Logics of conversation. Cambridge: Cambridge University Press.

Atkinson, Dwight (1999): Scientific Discourse in Sociohistorical Context: The Philosophical Transactions of the Royal Society of London 1675–1975. Mahwah, NJ.: Erlbaum.

Austin, John L. (1962): How to do things with words. The William James Lectures delivered at Harvard University in 1955. Edited by J. O. Urmson. Oxford: Oxford University Press

Austin, John L. (1970): Philosophical Papers. 2nd ed. Oxford: Oxford University Press.

Ausubel, David P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. Journal of Educational Psychology, 51, 267-272.

Bader, Anita (2018): Mailinglists als Format der digitalen Wissenschaftskommunikation. Eine linguistische Untersuchung. Gießen: Giessener Elektronische Bibliothek. http://geb.uni-giessen.de/geb/volltexte/2018/13523/.

Barnes, Julian (2011): The sense of an ending. London: Jonathan Cape.

Barnes, Julian (2019): The man in the red coat. London: Jonathan Cape.

Barthes, Roland (1942/1985): Notes sur André Gide et son *Journal*. In: Existences 27, juillet 1942, 7-18. Reprinted in: Bulletin des Amis d'André Gide 67, juillet 1985, 85-105.

Barthes, Roland (1977): Fragments d'un discours amoureux. Paris: Seuil.

Barzilay Regina/Lapata, Mirella (2008): Modeling local coherence: An entity-based approach. Computational Linguistics, 34(1), 1–34.

Bateman, John A. (2014): Text and image: a critical introduction to the visual-verbal divide. London [etc.]: Routledge.

Bateman, John A./Rondhuis, Klaas Jan (1997): Coherence relations: Towards a general specification. In: Discourse Processes 24, 3-49.

Bateman, John/Wildfeuer, Janina/Hiippala, Tuomo (2017): Multimodality. Foundations, research and analysis – a problem-oriented introduction. Berlin/Boston: De Gruyter.

Bazerman, Charles (1984): Modern Evolution of the experimental report in physics: spectroscopic articles in Physical Review, 1893-1980. In: Social Studies of Science 14(2), 163-196.

Bazerman, Charles (1987): Codifying the social science style: The APA Publication Manual as behavioristic rhetoric. In: Nelson, John S./Megill, Alan/McCloskey, Donald N. (eds.): The rhetoric of the human sciences. Language and argument in scholarship and public affairs. Madison: University of Wisconsin Press, 125-144.

Bazerman, Charles (1988): Shaping written knowledge. The genre and activity of the experimental article in science. Madison: University of Wisconsin Press.

Becher, Tony/Trowler, Paul R. (2001): Academic tribes and territories. 2nd ed. Ballmoor, Buckingham: The Society for Research into Higher Education & Open University Press.

Benamara, Farah/Taboada, Maite/Mathieu, Yannick (2017): Evaluative language beyond bags of words: Linguistic insights and computational applications. In: Computational Linguistics 43(1), 201-264.

Bender, Emily M./Gebru, Timnit/McMillan-Major, Angelina/Shmitchell, Shmargaret (2021): On the dangers of stochastic parrots: Can language models be too big? In. Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610-623. https://dl.acm.org/doi/10.1145/3442188.3445922.

Bender, Emily M./Koller, Alexander (2020): Climbing towards NLU: On meaning, form, and understanding in the age of data. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (July 5 – 10, 2020), 5185–5198.

Benz, Anton/Kühnlein, Peter (2008): Constraints in discourse – an introduction. In: Benz, Anton/Kühnlein, Peter (eds.): Constraints in discourse. Amsterdam/ Philadelphia: John Benjamins, 1-26.

Benz, Anton/Stede, Manfred (2012): Introduction. In: Benz, Anton/Stede, Manfred/ Kühnlein, Peter (eds.): Constraints in discourse 3: Representing and inferring discourse structure. Amsterdam/Philadelphia: John Benjamins, 1-16.

Bergmann, Jörg (1990): On the local sensitivity of conversation: In: Marková, Ivana/Foppa, Klaus (eds.): The dynamics of dialogue. New York: Harvester Wheatsheaf, 201-226.

Bergmann, Jörg (1993): Discreet indiscretions. The social organization of gossip. Abingdon, Oxfordshire: Taylor & Francis.

Geurts, Bart (2019): Communication as commitment sharing: speech acts, implicatures, common ground. In: Theoretical Linguistics 45(1-2), 1–30.

Biagioli, Mario (1993): Galileo courtier. The practice of science in the culture of absolutism. Chicago & London: The University of Chicago Press.

Biber, Douglas/Johansson, Stig/Leech, Geoffrey/Conrad, Susan/Finegan, Edward (1999): Longman Grammar of spoken and written English. Harlow: Pearson Education.

Bilal, Iman Munire/Wang, Bo/Liakata, Maria/Procter, Rob/Tsakalidis, Adam (2021): Evaluation of thematic coherence in microblogs. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics. Vol. 1, 6800–6814.

Blum, Joachim/Bucher, Hans-Jürgen (1998): Die Zeitung: Ein Multimedium. Textdesign – ein Gestaltungskonzept für Text, Bild und Grafik. Konstanz: UVK Medien.

Boy, Bettina/Bucher, Hans-Jürgen/Christ, Katharina (2020): Audiovisual science communication on TV and YouTube. How recipients understand and evaluate science videos. In: Frontiers in Communication, Vol. 5 (17 December 2020), 1-18. https://www.frontiersin.org/articles/10.3389/fcomm.2020.608620/full.

Bramhall, John (1655): A defence of true liberty from antecedent and extrinsecal necessity; being an answer to a late book of Mr. Thomas Hobbes of Malmesbury, entitled a treatise of liberty and necessity. London. In: The works of The Most Reverend Father in God, John Bramhall, D.D. Sometime Lord Archbishop of Armagh, Primate and Metropolitan of all Ireland. Vol. IV. Oxford: Parker 1844, 3-196.

Brandom, Robert B. (1994): Making it explicit. Reasoning, representing, and discursive commitment. Cambridge, Mass./London: Harvard University Press.

Brandt, Margareta/Rosengren, Inger (1992): Zur Illokutionsstruktur von Texten. In: Zeitschrift für Literaturwissenschaft und Linguistik 86, 9-51.

Breindl, Eva/Walter, Maik (2009): Der Ausdruck von Kausalität im Deutschen. Eine korpusbasierte Studie zum Zusammenspiel von Konnektoren, Kontextmerkmalen und Diskursrelationen. Mannheim: amades (IdS).

Brinker, Klaus/Antos, Gerd/Heinemann, Wolfgang/Sager, Sven Frederik (eds.) (2000/2001): Text- und Gesprächslinguistik. 2 vols. Berlin/New York: De Gruyter.

Brown, Tom B./Mann, Benjamin et al. (2020): Language models are few-shot learners. In: arXiv:2005.14165v4 [cs.CL] 22 Jul 2020. (https://arxiv.org/abs/2005.14165).

Bubenhofer, Noah/Scharloth, Joachim/Eugster, David (2015): Rhizome digital: Datengeleitete Methoden für alte und neue Fragestellungen in der Diskursanalyse. In: Zeitschrift für Diskursforschung 1. Sonderheft Diskurs, Interpretation, Hermeneutik, 144-172.

Bucher, Hans-Jürgen (1986): Pressekommunikation. Grundstrukturen einer öffentlichen Form der Kommunikation aus linguistischer Sicht. Tübingen: Niemeyer.

Bucher, Hans-Jürgen (1999): Die Zeitung als Hypertext. Verstehensprobleme und Gestaltungsprinzipien für Online-Zeitungen. In: Lobin, Hennig (Hg.): Text im digitalen Medium. Linguistische Aspekte von Textdesign, Texttechnologie und hypertext engineering. Opladen: Westdeutscher Verlag, 9-32.

Bucher, Hans-Jürgen (2001): Von der Verständlichkeit zur Usability. Rezeptionsbefunde zur Nutzung von Online-Medien. In: Osnabrücker Beiträge zur Sprachtheorie 63, 45-66.

Bucher, Hans-Jürgen (2017): Understanding multimodal meaning making: Theories of multimodality in the light of reception studies, In: Seizov, Ognyan/Wildfeuer Janina (eds.): New studies in Multimodality. Conceptual and methodological elaborations. London [etc.]: Bloomsbury, 91-123.

Bucher, Hans-Jürgen (2019): Politische Meinungsbildung in sozialen Medien? Interaktionsstrukturen in der Twitter-Kommunikation. In: Marx, Konstanze/Schmidt Axel (eds.): Interaktion und Medien. Interaktionslinguistische Zugänge zu medienvermittelter Kommunikation. Heidelberg: Winter, 287–318.

Bucher, Hans-Jürgen/Boy, Bettina/Christ, Katharina (2022): Audiovisuelle Wissenschaftskommunikation auf YouTube. Eine Rezeptionsstudie zur Vermittlungsleistung von Wissenschaftsvideos. Wiesbaden: Springer VS.

Bucher, Hans-Jürgen/Niemann, Philipp (2012): Visualizing science: The reception of PowerPoint presentations. In: Visual Communication 11, 283-306.

Burton, Robert (1964/1621): The anatomy of melancholy. In three Volumes. London: Dent. (Everyman's Library).

[Cambridge Grammar]: see Huddleston/Pullum (2002).

Carlson, Lynn/Marcu, Daniel (2001): Discourse tagging reference manual. Technical Report. Informations Science Institute, Marina del Rey, CA. ISI-TR-545.

Carston, Robyn (2009): The explicit/implicit distinction in pragmatics and the limits of explicit communication. In: International Review of Pragmatics 1, 35-62.

Chafe, Wallace L. (1976): Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In: Li, Charles N. (ed.): Subject and topic. New York/San Francisco/London: Academic Press, 27-55.

Chafe, Wallace L. (1979): The flow of thought and the flow of language. In: Givón, Talmy (ed.): Syntax and Semantics. Volume 12: Discourse and Syntax. San Diego etc.: Academic Press, 159-181.

Chafe, Wallace L. (1997): Polyphonic topic development. In: Givón, T. (ed.): Conversation: Cognitive, communicative and social perspectives. Amsterdam/ Philadelphia: John Benjamins, 41-53.

Clark, Herbert H. (1975). Bridging. In: Schank, Roger C./Nash-Webber, Bonnie L. (Hg.): Theoretical issues in natural language processing. New York: Association for Computing Machinery, 169-174.

Collins, Randall (1998): A global theory of intellectual change. Cambridge, Mass./London, England: Harvard University Press.

Crispin, Edmund (1958): The moving toyshop. Penguin Books. (Original 1946).

Dascal, Marcelo (1979): Conversational relevance. In: Margalit, Avishai (ed.): Meaning and use. Dordrecht [etc.]: Reidel, 153-174.

Dascal, Marcelo (2003): Interpretation and understanding. Amsterdam/Philadelphia: John Benjamins.

Dascal, Marcelo/Katriel, Tamar (1979): Digression: A study in conversational coherence. In: Petöfi, Janos S. (ed.) Text vs. Sentence. Continued. Hamburg: Buske, 76-95.

Dawkins, Richard (2003): A Devil's chaplain. Selected essays. London: Weidenfeld & Nicolson.

D'Costa, Krystal (2011): Smells from the past: A scent of a memory at the Fulton Fish Market. In: Goldman, Jason, G. (ed.): The open laboratory 2010. Chapel Hill, N.C.: Coturnix, 63-68.

Deleuze, Gilles/Guattari, Félix (1977): Rhizom. Berlin: Merve Verlag.

DeSalle, Rob/Yudell, Michael (2020): Welcome to the genome. A user's guide to the genetic past, present, and future. 2nd ed. Hoboken, N.J.: Wiley Blackwell.

Dorval, Bruce (ed.) (1990): Conversational organization and its development. Advances in discourse processes volume XXXVIII. Norwood, N.J.: Ablex Publishing Company.

Dummett, Michael (1973): Frege. Philosophy of language. London: Duckworth.

Dummett, Michael (1980): Twelve Tarot games. London: Duckworth.

Eemeren, Frans H. van/Grootendorst, Rob/Snoek Henkemans, Francisca (2002): Argumentation: Analysis, evaluation, presentation. Mahwah, NJ: Erlbaum.

Enfield, N.J./Levinson, Stephen C. (eds.): Roots of Human Sociality: Culture, cognition and interaction. Oxford/New York: Berg.

Ermakova, Vera (2015): Städte-, Länder- und Landschaftsporträts in gedruckten Zeitungen und Online-Angeboten. Grundlagen und empirische Untersuchungen zu Konstruktionsprinzipien einer multimodalen Kommunikationsform. Gießener Elektronische Bibliothek. Linguistische Untersuchungen Bd. 8. http://geb.uni-giessen.de/geb/volltexte/2015/11811/.

Everett, Daniel (2012): Language. The cultural tool. London: Profile Books.

Fabricius-Hansen, Cathrine (1999): Information packaging and translation. Aspects of translational sentence splitting (German – English/Norwegian). In: Doherty, Monika (Hg.): Sprachspezifische Aspekte der Informationsverteilung. Berlin: Akademie-Verlag, 175-214.

Fischer, Andreas (ed.): Repetition. Tübingen: Niemeyer.

Ford, Cecilia E./Thompson, Sandra A. (1996): Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns. In: Ochs, Elinor/Schegloff, Emanuel A./Thompson, Sandra A. (eds.): Interaction and Grammar. Cambridge: Cambridge University Press, 134-184.

Fox, Barbara A. (1986): Discourse structure and anaphora: written and conversational English. Cambridge: Cambridge University Press.

Friedman, Richard C./Downey, Jennifer I. (2008): Sexual orientation and psychodynamic psychotherapy. Sexual science and clinical practice. New York: Columbia University Press.

Fritz, Fiona (2022): The multifunctionality of a medieval hagiography. A historical case study of the Gesta et Passio and the making of the Danish Royal Saint Cnut (c. 1100). Stuttgart: Franz Steiner Verlag.

Fritz, Gerd (1982): Kohärenz. Grundfragen der linguistischen Kommunikationsanalyse. Tübingen: Narr.

Fritz, Gerd (1994): Grundlagen der Dialogorganisation. In: Fritz, Gerd/Hundsnurscher, Franz (eds.): Handbuch der Dialoganalyse. Tübingen: Niemeyer.

Fritz, Gerd (1997): Coreference in dialogue. In: Weigand, Edda (ed.): Dialogue analysis: Units, relations and strategies beyond the sentence. Tübingen: Niemeyer, 75-88.

Fritz, Gerd (1999): Coherence in hypertext. In: Bublitz, Wolfram/ Lenk, Uta/Ventola, Eija (eds.): Coherence in spoken and written discourse. Amsterdam: Benjamins, 221-234

Fritz, Gerd (2005): On answering accusations in controversies. In: Studies in Communication Sciences 5 [Special Issue: Argumentation in Dialogic Interaction], 151−162.

Fritz, Gerd (2008): Communication principles for controversies. A historical perspective. In: van Eemeren, Frans. H./Garssen, Bart (eds.): Controversy and confrontation: relating controversy analysis with argumentation theory. Amsterdam/ Philadelphia: John Benjamins, 109-124.

Fritz, Gerd (2011): Texttypen in wissenschaftlichen Blogs. Eine exemplarische Analyse am Beispiel des Language Log. In: Gloning, Thomas/Fritz, Gerd (eds.): Digitale Wissenschaftskommunikation. Formate und ihre Nutzung. Gießen: Gießener Elektronische Bibliothek, 205-285. http://geb.uni-giessen.de/geb/volltexte/ 2011/8227/.

Fritz, Gerd (2014): Texttheorie und textlinguistische Annotation von Korpora: Bemerkungen zur Rhetorical Structure Theory. In: Zeitschrift für Germanistische Linguistik 42, 163–195.

Fritz, Gerd (2017): Dynamische Texttheorie. 2nd ed. Gießen: Giessener Elektronische Bibliothek. http://geb.uni-giessen.de/geb/volltexte/2017/12601/.

Fritz, Gerd (2018a): The pragmatic organization of controversies. In: Fritz, Gerd/ Gloning, Thomas/Glüer, Juliane (eds.) Historical pragmatics of controversies. Case studies from 1600 to 1800. Amsterdam/Philadelphia: John Benjamins, 1-36.

Fritz, Gerd (2018b): Controversy and conversion. Friedrich Albert Carl Gren and the phlogiston controversy. In: Fritz, Gerd/Gloning, Thomas/Glüer, Juliane (eds.) Historical pragmatics of controversies. Case studies from 1600 to 1800. Amsterdam/Philadelphia: John Benjamins, 297-337.

Fritz, Gerd (2020a): Darstellungsformen in der historischen Semantik. Linguistische Untersuchungen. Bd. 14. Gießen: Giessen University Library Publications. http://geb.uni-giessen.de/geb/volltexte/2020/15084/.

Fritz, Gerd (2020b): Scholarly reviewing. In: Leßmöllmann, Annette/Dascal, Marcelo/Gloning, Thomas (eds.): Science communication. Boston/Berlin: De Gruyter, 289-310.

Fritz, Gerd (2020c): Scientific controversies. In: Leßmöllmann, Annette/Dascal, Marcelo/Gloning, Thomas (eds.): Science communication. Boston/Berlin: De Gruyter, 311-334.

Fritz, Gerd/Glüer, Juliane (2018): Reviews and responses. A controversy about the biblical canon. In: Fritz, Gerd/Gloning, Thomas/Glüer, Juliane: Historical pragmatics of controversies. Case studies from 1600 to 1800. Amsterdam/Philadelphia: John Benjamins, 253-296.

Fritz, Gerd/Hundsnurscher, Franz (1975): Sprechaktsequenzen. Überlegungen zur Vorwurf/Rechtfertigungs-Interaktion. In: Der Deutschunterricht 27, Heft 2, 81-103.

Gary, Marcus/Davis, Ernest (2020a): GPT-3, Bloviator: OpenAI's language generator has no idea what it's talking about. (https://www.technologyreview.com/ 2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion/ (August 22, 2020)).

Gary, Marcus/Davis, Ernest (2020b): Experiments testing GPT-3's ability at commonsense reasoning: results. https://cs.nyu.edu/~davise/papers/GPT3CompleteTests.html.

Geach, Peter Thomas (1968): Reference and generality. An examination of some medieval and modern theories. Emended edition. Ithaca/London: Cornell University Press.

Geach, Peter Thomas (1976): Reason and argument. Oxford: Blackwell.

Georgakopoulou, Alexandra (2007): Small stories, interaction and identities. Amsterdam/Philadelphia: John Benjamins.

George, Elizabeth (2004): Write away. One novelist's approach to fiction and the writing life. London: Hodder & Stoughton.

Gernsbacher, Morton Ann (1996): The structure-building framework: What it is, what it may also be, and why. In: Britton, Bruce K./Graesser, Arthur C. (eds.): Models of understanding text. Mahwah, N.J.: Lawrence Erlbaum, 289-311.

Gernsbacher, Morton Ann (1997): Coherence cues mapping during comprehension. In: Costermans, Jean/Fayol, Michel (eds.): Processing interclausal relationships in the production and comprehension of text. Hillsdale, NJ: Erlbaum, 3-21.

Gernsbacher, Morton Ann/Tallent, Kathleen A./Bolliger, Caroline N. (1999): Disordered discourse in schizophrenia described by the structure building framework. In: Discourse Studies 1(3), 355-372.

Gilbert, Stuart (1963): James Joyce's Ulysses. A study. London: Penguin.

Giora, Rachel (1985): Notes towards a theory of text coherence. In: Poetics Today, 6, 699-715.

Giora, Rachel (1997). Discourse coherence and theory of relevance: Stumbling blocks in search of a unified theory. In: Journal of Pragmatics 27, 17-34.

Giora, Rachel (1998). Discourse coherence is an independent notion: A reply to Deirdre Wilson's reply. In: Journal of Pragmatics 29, 75-86.

Giora, Rachel (2002): Masking one's themes. Irony and the politics of indirectness. In: Louwerse, Max/van Peer, Willie (eds.): Thematics. Interdisciplinary studies. Amsterdam/Philadelphia, 283-300.

Givón, Talmy (1983): Topic continuity in discourse: an introduction: In: Givón, Talmy (ed.): Topic continuity in discourse: A quantitative cross-language study. Amsterdam/Philadelphia: John Benjamins, 5-41.

Givón, T. (1993): Coherence in text, coherence in mind. In: Pragmatics & Cognition 1(2), 171–227.

Givón, T. (2005): Context as other minds. The pragmatics of sociality, cognition and communication. Amsterdam/Philadelphia: John Benjamins.

Givón, T. (2020): Coherence. Amsterdam/Philadelphia: John Benjamins.

[Glenn Gould] The Glenn Gould reader. Ed. by Tim Page. London/Boston: Faber and Faber 1987.

Gloning, Thomas (1996): Zur Vorgeschichte von Darstellungsformen und Textmerkmalen der ersten Zeitungen. In: Fritz, Gerd/Straßner, Erich (eds.): Die Sprache der ersten deutschen Wochenzeitungen im 17. Jahrhundert. Tübingen: Niemeyer, 196-258.

Gloning, Thomas (2008): „Man schlürft Schauspielkunst ...". Spielarten der Theaterkritik. In: Hagestedt, Lutz (ed.): Literatur als Lust. Begegnungen zwischen Poesie und Wissenschaft. Festschrift für Thomas Anz. München: belleville, 59-86.

Gloning, Thomas (2012): Textuelle Praktiken im Zwischenraum. Verfahren der thematischen Organisation in Rudolf Virchows Cellularpathologie (1858). In: Wirth, Uwe (Hg.): Bewegen im Zwischenraum. Berlin: Kulturverlag Kadmos, 337-358.

Gloning, Thomas (2015): Textkomposition und Multimodalität in Thurneyssers Buch über die Erdgewächse (1578). Eine Erkundung. In: Schuster, Britt-Marie/Dogaru, Dana Janetta (Hg.): Wirksame Rede im Frühneuhochdeutschen: Syntaktische und textstilistische Aspekte. Unter Mitarbeit von Arnika Lutz. Hildesheim/Zürich/ New York: Georg Olms Verlag, 177-211.

Gloning, Thomas (2020): Epistemic genres. In: Leßmöllmann, Annette/Dascal, Marcelo/Gloning, Thomas (eds.): Science communication. Boston/Berlin: De Gruyter, 209-233.

Goffman, Erving (1971): Relations in public. Microstudies of the public order. London: The Penguin Press.

Goffman, Erving (1981): Forms of talk. Oxford: Blackwell.

Goldman, Alvin I. (1970): A theory of human action. London etc.: Prentice-Hall.

Goutsos, Dionysis (1997): Modeling discourse topic: Sequential relations and strategies in expository texts. Norwood, New Jersey: Ablex.

Graesser, Arthur C./Britton, Bruce K. (1996): Five metaphors for text understanding. In: Britton, Bruce K./Graesser, Arthur C. (eds.): Models of understanding text. Mahwah, New Jersey: Erlbaum, 341-351.

Graesser, Arthur C./Millis, Keith K./Zwaan, Rolf A. (1997): Discourse comprehension. In: Annual Review of Psychology 48, 163-189.

Gredel, Eva (2017): Digital discourse analysis and Wikipedia: Bridging the gap between Foucauldian discourse analysis and digital conversation analysis. In: Journal of Pragmatics 115, 99–114.

Grice, Paul (1989): Studies in the way of words. Cambridge, Mass./London: Harvard University Press.

Gross, Alan G./Harmon, Joseph E./Reidy, Michael S. (2002): Communicating Science. The scientific article from the 17th century to the present. West Lafayette, Indiana: Parlor Press.

Grosz, Barbara J./Sidner, Candace L. (1986): Attention, intentions, and the structure of discourse. In: Computational Linguistics 12, 175-204.

Grosz, Barbara J./Joshi, Aravind K./Weinstein, Scott (1995): Centering: A framework for modelling the local coherence of discourse. In: Computational Linguistics 21, 203-225.

Halliday, M.A.K./Hasan, Ruqaiya (1976): Cohesion in English. London: Longman.

Hamblin, C.L. (1970): Fallacies. London: Methuen.

Handke, Peter (1990): Versuch über die Jukebox. Frankfurt am Main: Suhrkamp.

Harweg, Roland (1968): Pronomina und Textkonstitution. München: Wilhelm Fink Verlag.

Hearst, Marti A. (1997): TextTiling: Segmenting text into multi-paragraph subtopic passages. In: Computational Linguistics 23(1), 33-64.

Heim, Irene Roswitha (1982): The semantics of definite and indefinite noun phrases. Ph.D. Dissertation. University of Massachusetts, Amherst 1982.

Heringer, Hans J. (1978): Practical semantics. A study in the rules of speech and action. Berlin/New York: De Gruyter.

Hobbes, Thomas (1656): The questions concerning liberty, necessity, and chance, clearly stated and debated between Dr. Bramhall, Bishop of Derry, and Thomas Hobbes of Malmesbury. London. In: The English works of Thomas Hobbes of Malmesbury; now first collected and edited by Sir William Molesworth, Bart. Vol. V. London 1841, 1– 455. Reprint Aalen: Scientia 1962.

Hobbs, Jerry R. (1979): Coherence and coreference. In: Cognitive Science 3, 67-90.

Hobbs, Jerry R. (1990): Topic drift. In: Dorval, Bruce (ed.): Conversational organization and its development. Norwood, N.J.: Ablex Publishing Company, 3-22.

Holt, Elizabeth Jane (1991): Figures of speech: An exploration of the use of idiomatic phrases in conversation. A thesis submitted for the qualification of D. Phil, to The University of York, Department of Sociology. March, 1991. https://etheses.whiterose.ac.uk/4250/. (03.03.2022).

[Horace] Quintus Horatius Flaccus: Opera. Ed. by David R. Shackleton Bailey. Berlin/New York: De Gruyter 2008.

Hovy, Eduard H./Maier, Elisabeth (1995): Parsimonious or Profligate: How many and which discourse structure relations? Unpublished ms. https://www.isi.edu/natural-language/people/hovy/papers/93discproc.pdf. (26.04.2021).

Huddleston, Rodney/Pullum, Geoffrey K. (2002): The Cambridge Grammar of the English Language. Cambridge: Cambridge University Press.

Hull, David L. (1988): Science as a process. An evolutionary account of the social and conceptual development of science. Chicago/London: University of Chicago Press.

Hundsnurscher, Franz (1981): On insisting. In: Parret, Herman/Sbisà, Marina/ Verschueren, Jef (eds.): Possibilities and limitations of pragmatics. Amsterdam: John Benjamins, 343-357.

Hundsnurscher, Franz (1997): Repetition reconsidered. In: Weigand, Edda (ed.): Dialogue analysis. Units, relations and strategies beyond the sentence. Tübingen: Niemeyer, 107-119

Hyland, Ken (1998): Hedging in scientific research articles. Amsterdam: John Benjamins

Hyland, Ken (2004): Disciplinary discourses. Social interactions in academic writing. Ann Arbor: University of Michigan Press.

Hyland, Ken/Diani, Giuliana (eds.) (2009): Academic evaluation: Review genres in university settings. Basingstoke: Palgrave Macmillan.

Hyland, Ken/Jiang, Feng (Kevin) (2020): Text-organizing metadiscourse. Tracking changes in rhetorical persuasion. In: Journal of Historical Pragmatics 21, 137-164.

[ICMJE Recommendations 2019] International Committee of Medical Journal Editors (2019): Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals (Updated December 2019). http://www.icmje.org/.

Jackson, Philip (2020): Understanding understanding and ambiguity in natural language. In: Procedia Computer Science 169, 209-225.

James, P.D. (2009): Talking about detective fiction. London: Faber and Faber.

Jaszczolt, Katarzyna M. (2018): Defaults in Semantics and Pragmatics. In: Edward N. Zalta (ed.): The Stanford Encyclopedia of Philosophy (Summer 2018 Edition). https://plato.stanford.edu/archives/sum2018/entries/defaults-semantics-pragmatics/.

Jefferson, Gail (1972): Side sequences. In: Sudnow, David (ed.): Studies in social interaction. New York: The Free Press, 294-338.

Jefferson, Gail (1984): On stepwise transition from talk about a trouble to inappropriately next-positioned matters. In Atkinson, J. Maxwell/Heritage, John (eds.): Structures of social action: Studies of conversation analysis. Cambridge: Cambridge University Press, 191-222.

Johnson, Ralph (2011): The coherence of Hamblin's *Fallacies*. In: Informal Logic 31, 305-317.

Jolley, Nicholas (1984): Leibniz and Locke: A study of the New Essays on Human Understanding. Oxford: Clarendon Press.

Jones, Peter (2007): Reading Ovid: Stories from the Metamorphoses. Cambridge: Cambridge University Press.

Jung, Matthias (1994): Öffentlichkeit und Sprachwandel. Zur Geschichte des Diskurses über die Atomenergie. Opladen: Westdeutscher Verlag.

Kaltwasser, Dennis (2019): Forenkommunikation in Onlinezeitungen. Pressekommunikation im medialen Wandel. Gießen: Giessen University Library Publications. http://geb.uni-giessen.de/geb/volltexte/2019/14812/.

Karamanis, Nikiforos (2004): Entity coherence for descriptive text structuring. PhD dissertation. University of Edinburgh.

Karamanis, Nikiforos/Poesio, Massimo/Mellish, Chris/Oberlander, Jon (2004): Evaluating Centering-based metrics of coherence for text structuring using a reliably annotated corpus. In: ACL '04: Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics, 391-398.(https://dl.acm.org/doi/10.3115/1218955.1219005.

Karttunen, Lauri (1976): Discourse referents. In: McCawley, James D. (ed.): Syntax and Semantics. Vol. 7. New York: Academic Press, 363-386.

Kasher, Asa (1987): Justification of speech, acts and speech acts. In: Le Pore, Ernest (ed.): New directions in semantics. London: Academic Press, 281-303.

Katz, David/Katz, Rosa (1928): Gespräche mit Kindern. Untersuchungen zur Sozialpsychologie und Pädagogik. Berlin: Julius Springer.

Keenan, Elinor O./Klein, Ewan (1975): Coherency in children's discourse. In: Journal of Psycholinguistic Research 4(4), 365-380.

Kehler, Andrew (2019): Cohesion and coherence. In: Portner, Paul/Maienborn, Claudia/Heusinger, Klaus von (eds.): Semantics. Sentence and information structure. Berlin/Boston: De Gruyter/Mouton, 450-480.

Kehler, Andrew/Kertz, Laura/Rohde, Hannah/Elman, Jeffrey, L. (2008): Coherence and coreference revisited. In: Journal of Semantics 25, 1-44.

Kelly, Thomas (2016): Evidence. In: Zalta, Edward N. (ed.): The Stanford Encyclopedia of Philosophy (Winter 2016 Edition), (https://plato.stanford.edu/archives/win2016/entries/evidence/).

Kendrick, Robin H./ Brown, Penelope/Dingemanse, Mark/Floyd, Simeon/Gipper, Sonja/Hayano, Kaoru/Hoey, Elliott/ Hoymann, Gertie/Manrique, Elizabeth/Rossi, Giovanni/Levinson, Stephen C. (2020): Sequence organization: A universal infrastructure for social action. In: Journal of Pragmatics 168, 119-138.

Kintsch, Walter (1998): Comprehension. A paradigm for cognition. Cambridge: Cambridge University Press.

Knoblauch, Hubert/Luckmann, Thomas (2004): Genre analysis. In: Flick, Uwe/von Kardoff, Ernst/Steinke, Ines (ed.): A companion to Qualitative Research. London etc.: Sage Publications, 305-307.

Knott, Alistair (1996): A Data-Driven Methodology for Motivating a Set of Coherence Relations. PhD thesis, Department of Artificial Intelligence, University of Edinburgh.

Knott, Alistair/Sanders, Ted (1998): The classification of coherence relations and their linguistic markers: An exploration of two languages. In: Journal of Pragmatics 30, 135-175.

Kopak, Richard W. (1999): Functional link typing in hypertext. In: ACM Computing SurveysVolume 31, Issue 4es, 16ff. (https://doi.org/10.1145/345966.346005).

Kress, Gunther/van Leeuwen, Theo (2001): Multimodal discourse. The modes and media of contemporary communication. London: Arnold.

Krifka, Manfred (2006): Basic notions of information structure. In: Féry, Caroline/ Fanselow, Gisbert/Krifka, Manfred (eds.): Interdisciplinary studies on informationstructure 6. Potsdam, 13-55. https://amor.cms.huberlin.de/~h2816i3x/ Publications/Krifka_InformationStructure.pdf.

Krifka, Manfred (2019): Commitments and beyond. In: Theoretical Linguistics 45(1-2), 73-91.

Kripke, Saul (1972): Naming and necessity. In: Davidson, Donald/Harman, Gilbert (eds.): Semantics of natural language. Dordrecht/Boston: Reidel, 253-355.

Kripke, Saul A. (2011): Vacuous names and fictional entities. In: Kripke, Saul A.: Philosophical troubles. Collected papers. Vol I. Oxford: Oxford University Press, 52-74.

Labov, William (1972): Ritual insults. In: Sudnow, David (ed.): Studies in social interaction. New York: The Free Press/London: Collier-Macmillan, 120-169.

Labov, William/Waletzky, Joshua (1967): Narrative analysis: Oral versions of personal experience. In: Helm, John (ed.): Essays on the verbal and visual arts. Seattle/London: University of Washington Press, 12-44.

Landow, George P. (1991): The rhetoric of hypermedia: some rules for authors. In: Delany, Paul/Landow, George P. (eds.): Hypermedia and Literary Studies. Cambridge, Mass.: MIT Press, 81-104.

Landow, George P. (2006): Hypertext 3.0. Critical theory and new media in an era of globalization. Baltimore: The Johns Hopkins University Press.

Leibniz, Gottfried Wilhelm (1996): New essays on human understanding. Translated and edited by Peter Remnant and Jonathan Bennett. Cambridge: Cambridge University Press.

Levinson, Stephen C. (2000): Presumptive meanings. Cambridge, Mass./London: The MIT Press.

Levinson, Stephen C. (2016): Turn-taking in human communication – Origins and implications for language processing. In: Trends in Cognitive Sciences 20(1), 6-14.

Lewis, David K. (1968): Convention: A philosophical study. Cambridge, Mass.: Harvard University Press.

Lewis, David (1979): Scorekeeping in a language game. In: Journal of Philosophical Logic 8, 339-359.

Libermann, Mark/Pullum, Geoffrey K. (2006): Far from the madding gerund and other dispatches from LANGUAGE LOG. Wilsonville, Oregon: William, James & Co.

Lin, Chin-Yew/Hovy, Eduard (1997): Identifying Topics by Position. In: Proceedings of the 5th Conference on Applied Natural Language Processing (ANLP97). Washington, D.C., March 1997, 283-290.

Lin, Chin-Yew/Hovy, Eduard (2000): The Automated Acquisition of Topic Signatures for Text summarization. In: The 18th International Conference on Computational Linguistics (COLING 2000). Vol. 1, 495-501. https://aclanthology.org/C00-1072.

Linell, Per (1998): Approaching dialogue. Talk, interaction and context in dialogical perspectives. Amsterdam/Philadelphia: John Benjamins.

Linnik, Anastasia/Bastiaanse, Roelien/Stede, Manfred/Khudyakova, Mariya (2022): Linguistic mechanisms of coherence in aphasic and non-aphasic discourse. In: Aphasiology 36(2), 123-146.

Longacre, Robert E. (1979): The paragraph as a grammatical unit. In: Givón, Talmy (ed.): Syntax and Semantics. Volume 12: Discourse and Syntax. San Diego etc.: Academic Press, 115-134.

[Longman Grammar]: see Biber et al. (1999).

Luckmann, Thomas (1986): Grundformen der gesellschaftlichen Vermittlung des Wissens: Kommunikative Gattungen. In: Neidhardt, Friedhelm/ Lepsius, M. Rainer/Weiß, Johannes (eds.): Kultur und Gesellschaft. (Sonderheft 27 der Kölner Zeitschrift für Soziologie und Sozialpsychologie), 191−211.

McCoy, R. Thomas/Pavlick, Ellie/ Linzen, Tal (2019): Right for the wrong reasons: Diagnosing syntactic heuristics in Natural Language Inference. In: In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA (USA), 3428–3448.

Mackie, Penelope (2014): Counterfactuals and the fixity of the past. In: Philosophical Studies168, 397-415.

Mann, William C./Thompson, Sandra A. (1988): Rhetorical structure theory: Toward a functional theory of text organization. In: Text 8, 243-281.

Marcu, Daniel (2000): The rhetorical parsing of unrestricted texts: A surface-based approach. In: Computational Linguistics 26(3), 395-448.

Marcu, Daniel/Romera, Magdalena/Amorrortu, Estibaliz (1999): Experiments in constructing a corpus of discourse trees: Problems, annotation choices, issues. In: Towards standards and tools for discourse tagging. https://aclanthology.org/W99-0307.pdf, 48-57.

Marneffe, Marie-Catherine de/Rafferty, Anna N./Manning Christopher D. (2008): Finding contradictions in text. In: Proceedings of ACL-08: HLT, 1039–1047. https://nlp.stanford.edu/pubs/contradiction-acl08.pdf

Meier, Mischa (2019): Geschichte der Völkerwanderung. Europa, Asien und Afrika vom 3. bis zum 8. Jahrhundert n.Chr. München: C.H. Beck Verlag.

Merton, Robert K. (1965): On the shoulders of giants. A Shandean postscript. New York/London: Harcourt Brace Jovanovich.

Millis, Keith K./King, Anne/Kim, Hyun-Jeong J. (2000): Updating situation models from descriptive texts. A test of the situational operator model. In: Discourse Processes 30, 201-236.

Moore, Johanna/Pollack, Martha (1992): A problem for RST: The need for multi-level discourse analysis. In: Computational Linguistics 18(4), 537-544.

Motta-Roth, Désirée (1998): Discourse analysis and academic book reviews: a study of text and disciplinary cultures. In: Fortanet, Inmaculada/Posteguillo, Santiago/Palmer, Juan Carlos/Coll, Juan Francisco (eds.): Genre Studies in English for Academic Purposes. Castelló, Spain: Universitat Jaume I, 29-58.

Muckenhaupt, Manfred (1978): Lernziel sprachliches Handeln. Beispiele für einen kommunikativen Sprachunterricht in der Sekundarstufe I. München: Ehrenwirth.

Muckenhaupt, Manfred (1986): Text und Bild. Grundfragen der Beschreibung von Text-Bild-Kommunikationen aus sprachwissenschaftlicher Sicht. Tübingen: Narr.

Nadler, Steven (1999): Spinoza. A life. Cambridge: Cambridge University Press.

Nie, Yixin/Williams, Adina/Dinan, Emily/Bansal, Mohit/Weston, Jason/Kiela, Douwe (2019): Adversarial NLI: A new benchmark for natural language understanding. In: arXiv:1910.14599 [cs.CL] (31 Oct 2019). https://arxiv.org/abs/1910.14599. (01.03.22)

Paperno, Denis/Kruszewski, German/Lazaridou, Angeliki et al. (2016): The LAMBADA dataset: Word prediction requiring a broad discourse context. In: arXiv:1606.06031v1 [cs.CL] (20 Jun 2016). https://arxiv.org/pdf/1606.06031.pdf. (01.03.22).

Partee, Barbara Hall (1971): On the requirement that transformations preserve meaning. In: Fillmore, Charles J./Langendoen, Terence D. (eds.): Studies in linguistic semantics. New York/Chicago [etc.]: Holt, Rinehart and Winston, 1-21.

Peldszus Andreas/Stede, Manfred (2013): From argument diagrams to argumentation mining in texts: a survey. In: International Journal of Cognitive Informatics and Natural Intelligence (IJCINI) 7(1), 1-31.

[PNAS template] Proceedings of the National Academy of Sciences (PNAS): Template for preparing your research report submission to PNAS. https://www.pnas.org/author-center.

Poesio, Massimo/Stevenson, Rosemary/Di Eugenio, Barbara/Hitzeman, Janet (2004): Centering: a parametric theory and its instantiations. In: Computational Linguistics 30, 309-363.

Prasad, Rashmi/Dinesh, Nikhil/Lee, Alan/Miltsakaki, Eleni/Robaldo, Livio/Joshi, Aravind/Webber, Bonnie (2008): The Penn Discourse TreeBank 2.0. In: Proceedings of the 6th International Conference on Language Resources and Evaluation, 2961-2968.

Pulvermüller, Friedemann (1990): Aphasische Kommunikation. Grundfragen ihrer Analyse und Therapie. Tübingen: Gunter Narr.

Qiang, Jipeng/Qian, Zhenyu/Li, Yun/Yuan, Yunhao/Wu, Xindong (2022): Short text topic modeling techniques, applications, and performance: A survey. In: IEEE Transactions on Knowledge and Data Engineering 34(3), 1427-1445.

Quine, Willard van Orman (1960): Word and object. Cambridge, Mass.: The M.I.T Press.

Quintilianus, Marcus Fabius (1995): Institutionis oratoriae libri XII, ed. Helmut Rahn. Darmstadt: Wissenschaftliche Buchgesellschaft.

Riou, Marine (2017): Transitioning to a new topic in American English conversation: A multi-level and mixed-methods account. In: Journal of Pragmatics 177, 88-105.

Rösner, Dietmar/Stede, Manfred (1992): Customizing RST for the automatic production of technical manuals. In: Dale, Robert/Hovy, Eduard et al. (eds.): Aspects of automated natural language generation. Berlin/Heidelberg/New York: Springer-Verlag, 199-214.

Ross, John Robert (1967): Constraints on variables in syntax. PhD dissertation. Massachusetts Institute of Technology.

Sacks, Harvey (1992): Lectures on conversation. 2 vols. Oxford/Cambridge, Mass.: Blackwell.

Sacks, Harvey/ Schegloff, Emanuel A./Jefferson, Gail (1974): A simplest systematics for the organization of turn-taking for conversation. In: Language 50(4), 696-735.

Schatzki, Theodore R. (1996): Social practices. A Wittgensteinian approach to human activity and the social. Cambridge: Cambridge University Press.

Schegloff, Emanuel A. (1972): Notes on a conversational practice: Formulating place. In: Sudnow, David (ed.): Studies in social interaction. New York: The Free Press, 75-119

Schegloff, Emanuel A. (1999): On the organization of sequences as a source of "coherence" in talk-in-interaction. In: Dorval, Bruce (ed.): Conversational organization and its development. Norwood, New Jersey: Ablex, 51-77.

Schegloff, Emanuel A. (2007): Sequence organization in interaction: A primer in Conversation Analysis. Cambridge: Cambridge University Press.

Schegloff, Emanuel A./Sacks, Harvey (1973): Opening up closings. In: Semiotica 8(4), 289-327.

[Schiller/Goethe correspondence] Schiller/Goethe Briefwechsel. Hg. von Emil Staiger. Revidierte Neuausgabe von Hans-Georg Dewitz. Frankfurt am Main: Insel Verlag 2005.

Schleiermacher, Friedrich (1998): Hermeneutics and criticism and other writings. Edited by Andrew Bowie. Cambridge: Cambridge University Press.

Schröder, Thomas (1995): Die ersten Zeitungen. Textgestaltung und Nachrichtenauswahl. Tübingen: Narr.

Schröder, Thomas (2003): Die Handlungsstruktur von Texten. Ein integrativer Beitrag zur Texttheorie. Tübingen: Narr.

Searle, John R. (1969): Speech acts. An essay in the philosophy of language. Cambridge: Cambridge University Press.

Searle, John R. (1979): The logical status of fictional discourse. In: Searle, John R.: Expression and meaning. Studies in the theory of speech acts. Cambridge: Cambridge University Press, 58-75.

Searle, John R. (1983): Intentionality. An essay in the philosophy of mind. Cambridge: Cambridge University Press.

Shwayder, D.S. (1965): The stratification of behaviour. A system of definitions propounded and defended. London: Routledge & Kegan Paul.

Somerville, Margaret (2006): The ethical imagination: Journeys of the human spirit. Toronto, Ontario: House of Anansi Press.

Stalnaker, Robert C. (1998): On the representation of context. Journal of Logic, Language, and Information 7, 3-19.

Stalnaker, Robert C. (2002): Common ground. In: Linguistics and Philosophy 25, 701-721.

Stark, Heather A. (1988) What do paragraph markings do? Discourse Processes, 11(3), 275-303.

Stede, Manfred (2007): Korpusgestützte Textanalyse. Grundzüge der Ebenenorientierten Textlinguistik. Tübingen: Narr. 2nd revised ed. Tübingen: Narr (2018).

Stede, Manfred (2012): Discourse processing. San Rafael, California: Morgan & Claypool Publishers.

Stede, Manfred/Peldszus, Andreas (2012): The role of illocutionary status in the usage conditions of causal connectives and in coherence relations. In: Journal of Pragmatics 44, 214-229.

Stegert, Gernot (1993): Filme rezensieren in Presse, Radio und Fernsehen. München: TR-Verlagsunion.

Sterne, Laurence (1962): The life and opinions of Tristram Shandy, Gentleman. New York: The New American Library. (Original 1759-1767).

Strawson, Peter F. (1952): Introduction to logical theory. London: Methuen.

Strawson, Peter F. (1959): Individuals. An essay in descriptive metaphysics. London: Methuen.

Strawson, Peter F. (1964/1971): Identifying reference and truth-values. In: Theoria 30, 1964, 96-118. Reprinted in: Strawson, Peter F. (1971): Logico-linguistic papers. London: Methuen, 75-95.

Strawson, Peter F. (1974): Subject and predicate in logic and grammar. London: Methuen.

Subba, Rajen/Di Eugenio, Barbara (2009): An effective discourse parser that uses rich linguistic information: In: Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the ACL. Boulder, Colorado, 566-574.

Swales, John M. (1990): Genre analysis. English in academic and research settings. Cambridge: Cambridge University Press.

Swales, John M. (2004): Research genres. Explorations and applications. Cambridge: Cambridge University Press.

Taboada, Maite (2006): Discourse markers as signals (or not) of rhetorical relations. In: Journal of Pragmatics 38(4), 567–592

Taboada, Maite/Brooke, Julian/Stede, Manfred (2009): Genre-based paragraph classification for sentiment analysis. Proceedings of SIGDIAL 2009: the 10th Annual Meeting of the Special Interest Group in Discourse and Dialogue. Queen Mary University of London, September 2009. Association for Computational Linguistics, 62–70.

Taboada, Maite/Mann, William C. (2006): Rhetorical structure theory: Looking back and moving ahead. In: Discourse Studies 8, 423–459.

Tamkin, Alex/Brundage, Miles/Clark, Jack/Ganguli, Deep (2021): Understanding the capabilities, limitations, and societal impact of large language models. In: arXiv:2102.02503v1 [cs.CL] 4 Feb 2021. (https://arxiv.org/abs/2102.02503).

Tannen, Deborah (2007): Talking voices. Repetition, dialogue, and imagery in conversational discourse. 2nd ed. Cambridge: Cambridge University Press.

Thagard, Paul (2000): Coherence in thought and action. Cambridge, Mass.: The MIT Press.

Thagard, Paul/Verbeurgt, Karsten (1998): Coherence as constraint satisfaction. In: Cognitive Science 22(1), 1-24.

Thomason, Richmond. H. (1990): Accomodation, meaning, and implicature: Interdisciplinary foundations for Pragmatics. In: Cohen, Philip/Morgan, Jerry/Pollack, Martha E. (eds.): Intentions in communication. Cambridge, Mass./London: MIT Press, 325-363.

Tomasello, Michael (2006): Why don't apes point? In: Enfield, N.J./Levinson, Stephen C. (eds.): Roots of Human Sociality: Culture, cognition and interaction. Oxford/New York: Berg, 506-524.

Tomasello, Michael (2008): Origins of human communication. Cambridge, Mass./New York: The MIT Press.

Toulmin, Stephen E. (1958): The uses of argument. Cambridge: Cambridge University Press.

Toulmin, Stephen E./Baier, Kurt (1952): On describing. In: Mind 61, 13-38.

Ullmann-Margalit, Edna (1978): Invisible-hand explanations. In: Synthese 39, 263-291.

Valentin, Karl (1969): Sturzflüge im Zuschauerraum. Der gesammelten Werke anderer Teil. Herausgegeben von Michael Schulte. München: Piper Verlag.

Vanderveken, Daniel (1990): Meaning and speech acts. Vol 1: Principles of language use. Cambridge: Cambridge University Press.

Van der Vliet, Nynke/Redeker, Gisela (2014): Explicit and implicit coherence relations in Dutch texts. In: Gruber, Helmut/Redeker, Gisela (eds.): The pragmatics of discourse coherence. Theories and applications. Amsterdam/Philadelphia: John Benjamins, 23-52.

Walton, Douglas (2007): Dialog theory for critical argumentation. Amsterdam/Philadelphia: John Benjamins.

Weiss, Edmond H. (1991): How to write usable user documentation. 2nd ed. Phoenix, Az.: Oryx Press.

Wilson, Deirdre (1998): Discourse, coherence and relevance: a reply to Rachel Giora. In: Journal of Pragmatics 29, 57-74.

Wittgenstein, Ludwig (1921): Tractatus Logico-Philosophicus (Logisch-Philosophische Abhandlung). In: Annalen der Naturphilosophie 44, 185-262.

Wittgenstein, Ludwig (1958): Philosophical Investigations. Ed. by G.E.M. Anscombe and R. Rhees, translated by G.E.M. Anscombe. Oxford: Blackwell.

Wittgenstein, Ludwig (1969): On certainty. (Über Gewißheit). Ed. by G.E.M. Anscombe and Georg Henrik von Wright. Translated by Denis Paul and G.E.M. Anscombe. Oxford: Basil Blackwell.

Wolańska-Köller, Anna (2010): Funktionaler Textaufbau und sprachliche Mittel in Kochrezepten des 19. und frühen 20. Jahrhunderts. Stuttgart: *ibidem*- Verlag.

Wolf, Florian/Gibson, Edward (2006): Coherence in natural language. Data structures and applications. Cambridge, Mass.: The MIT Press.

Woolf, Virginia (1964): Mrs Dalloway. Harmondsworth, England: Penguin Books. (Original 1925).

Wright, Georg Henrik von (1963): The varieties of goodness. London: Routledge & Kegan Paul.

Wu, Jianguo (2011): Improving the writing of research papers: IMRAD and beyond. In: Landscape Ecology 26, 1345–1349.

Xu, Peng/Saghir, Hamidreza/Sung Kang, Jin et al. (2019): A cross-domain transferable neural coherence model. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 678–687. (Association for Computational Linguistics). https://aclanthology.org/P19-1067/.

Ziem, Alexander (2014): Frames of understanding in text and discourse. Theoretical foundations and descriptive applications. Amsterdam/Philadelphia: John Benjamins.

Zwaan, Rolf A. (1996): Toward a model of literary comprehension. In: Britton, Bruce, K./Graesser, Arthur C. (eds.): Models of understanding text. Mahwah, N.J.: Erlbaum, 241-255.

Texts and (contributions to) conversations mostly show a kind of orderliness and serial connectedness that makes them comprehensible and useful. Accordingly, a theory of coherence has to account for the resources and strategies used by speakers/writers in achieving this orderliness and connectedness and by hearers/readers in understanding it. The present book approaches this task from an action-theoretic perspective, focusing in particular on the dynamics of act sequencing and modelling the step-by-step solution of communicative tasks in the course of text production and dialogical communication. Individual chapters (and sections) deal with the following topics, among others: the basic building blocks of discourse, resources of coherence and their use, speech act sequencing, coherence and coreference, coherence and topic management, coherence and text types, coherence and understanding, coherence and intertextuality, types of connectivity in text/image configurations, coherence and Artificial Intelligence, topics as resource and an achievement in dialogue.

9 783347 709409

Linguistische Untersuchungen 16

Herausgegeben von Gerd Fritz, Thomas Gloning und Dennis Kaltwasser