

## **Cognitive Discourse Analysis: accessing cognitive representations and processes through language data\***

THORA TENBRINK

*School of Linguistics and English Language, Bangor University, Wales, UK*

*(Received 18 October 2013 – Revised 22 January 2014 – Accepted 05 March 2014 –*

*First published online 14 July 2014)*

### ABSTRACT

This paper offers the first general introduction to CODA (Cognitive Discourse Analysis), a methodology for analyzing verbal protocols and other types of unconstrained language use, as a resource for researchers interested in mental representations and high-level cognitive processes. CODA can be used to investigate verbalizations of perceived scenes and events, spatio-temporal concepts, complex cognitive processes such as problem-solving and cognitive strategies and heuristics, and other concepts that are accessible for verbalization. CODA builds on and extends relevant established methodologies such as cognitive linguistic perspectives, verbal protocol analysis in cognitive psychology and interdisciplinary content analysis, linguistic discourse analysis, and psycholinguistic experimentation.

**KEYWORDS:** applied cognitive linguistics, verbalization of thought, conceptualization, problem solving, discourse analysis

### 1. Introduction

Language use is based on thoughts. Whatever we say must have gone through our minds, in one way or other, superficially or in depth. Thoughts can (to some extent at least) be put into words; people often ask “What are you thinking?”, and they expect a meaningful answer to follow. They may not

---

[\*] I feel privileged for the many opportunities of collaborating with diverse experts across disciplines. My sincere thanks go to the numerous project partners and collaborators in projects and publications mentioned throughout this paper, especially to Holly Taylor for intense collaboration over the past years, and to Vyv Evans for support and inspiring discussion. I am also grateful for diverse comments from many people on earlier versions of this paper, most prominently Michel Denis, Vivien Mast, Dan Montello, Holly Taylor, and Matthew Walsh. Address for correspondence: Thora Tenbrink, School of Linguistics & English Language, Room 306 Linguistics, Bangor University, 39 College Road, Bangor, Gwynedd, LL57 2DG, UK. tel: +44 1248 382263; fax: +44 1248 383267; e-mail: t.tenbrink@bangor.ac.uk

even realize that the answer will be indirect and communicated through a medium, typically LANGUAGE. Nevertheless there is no direct way of accessing thoughts, and the language people use to express them cannot be equated with their thoughts. The relationship between language and thought is not simple, but undoubtedly it is systematic (Miller, 1951). To the extent that systematic principles and patterns can be identified, they can be exploited for accessing what goes on in people's minds.

Since researchers interested in human thought and behavior frequently aim to access cognition, language is a widely used medium across various research purposes and procedural steps. This starts with (mostly spoken, sometimes written) discussions among researchers when first designing the procedure, is carried further through task instructions that are conveyed verbally in most cases, and may further involve behavioral responses given through language, or direct questions during task performance. Centrally language-based methods include verbal protocols (such as think-aloud data and retrospective reports), interviews, and informal discussions used for inspiration. Altogether, there are many ways of gaining insights through language. The relevant appearances of language are variously treated as data (to be analyzed according to specific features) or as medium (which, in itself, is not particularly interesting), analyzed ad hoc and intuitively, or remain altogether unmentioned in publications if authors feel that they played no appreciable part in the process (in spite of having served as considerable resources for inspiration). Dealing with language seems unproblematic; in a sense all of us are experts in the interpretation of this medium, or feel we are, since we all use it every day. However, in spite of the ubiquity of language in behavioral research, language is actually rarely treated from an expert point of view – i.e., analyzed in a rigorous way based on linguistic background knowledge. Thus, we use language as a medium and data resource to learn about thought – but to what extent do we know what we are doing, and how can we deal with this form of representation systematically rather than intuitively?

Let us consider some examples of language use. Imagine a person describing a visual scene, like a traffic situation on the road. Will they focus on the trees, the cars, the pedestrians, or the gray sky? Most centrally this will depend on relevance for the discourse task at hand (Sperber & Wilson, 1986). Guided by relevance, the speaker's linguistic choices will necessarily reflect their conceptualization of the scene in systematic ways. Their attentional focus determines the choice of objects and persons described. The gray sky will only be mentioned if the weather conditions are consciously noted by the speaker, due to perceived relevance for the current discourse, emotional affect, or for other conceptually anchored reasons. Upon closer analysis, the information structure of the speaker's description reveals which aspects are represented as new or taken for granted, and which are foregrounded

or remain implicit. Consider the following sentences, which constitute fundamentally different references to (possibly) the same scene at the time of speaking:

- (1) The blue car is parked in front of a tree.
- (2) Did you see how this idiot almost crashed into the tree?

Unlike (1), (2) is directed at an addressee, prominent in the speaker's mind. It contains a range of affective evaluations, reflected by terms such as *idiot* and *crashed*, and supported by the invitation (conveyed as a question *Did you see?*) to share the speaker's perception. The car remains implicit in this utterance but is inferable from the motion situation evoked by *crashed into the tree*; the color does not appear in the description and therefore does not matter to the speaker at this moment. In (1), the car is referred to as the given starting point for a description of its location, while the tree is introduced as new using the indefinite article; (2) presupposes the tree's existence (and accessibility to the addressee) as signaled by the definite article. Most strikingly, however, the second sentence reflects a conceptualization of a DYNAMIC scene (conveyed by the form of the verb), namely an event preceding the current view of the scene. In contrast, the first describes the current view as a STATIC scene – one that may be the visually available result of a dynamic procedure, though not the one described in (2): the speaker assumes that the car had merely been *parked*. Altogether, although both sentences may refer to the same visual scene, the speakers' linguistic choices convey their fundamentally different perceptions and conceptualizations in multiple ways. Here, a plausible explanation is that the second speaker may have had access to a more extended portion of the antecedent motion event, leading to the affective evaluation and the conceptual focus on the dynamic aspect.

In these and many other ways, language use reflects crucial aspects about the speakers' concepts, mediated by their understanding of the communicative situation, at any given moment. This provides a good pathway to access cognition, given the necessary expertise about relevant features of language. Features of a linguistic utterance that pertain to cognition are revealing about a speaker's thoughts and cognitive processes, and can thus inform cognitive science directly. Features of a linguistic utterance that pertain to communication are crucial for a wide range of applications within cognitive science, such as human–robot and human–computer interaction, automatically generated user support, intuitive assistance systems, and so forth.

The central idea in the methodological framework presented here, Cognitive Discourse Analysis (CODA), is to use unconstrained natural language elicited in purposefully controlled situations as a data source; ideally combined with other modalities or representations of cognitive processes. Across various

recent projects,<sup>1</sup> research questions about human cognition have been addressed using this methodological framework. One overarching aim in this research has been to accumulate insights about how and to what extent language analysis can support cognitive science research. Results include, for instance, ways in which speakers switch flexibly between conceptual domains, the flexibility and range of problem-solving strategies within and across speakers, and the impact of situation and discourse context on linguistic representation. Various examples will be given below to illustrate the methodology.

This paper provides the first general introduction to CODA as a tool for analyzing the language that speakers use to express thought. Relevant research questions broadly fall into two areas: *MENTAL REPRESENTATION* (the conceptualization of complex scenes, event perception, and the like), and *COMPLEX COGNITIVE PROCESSES* (such as problem-solving or decision-making). Both of these relate to and enhance well-established research traditions in distinct ways. With respect to mental representation, CODA addresses the conceptualization of perceived situations and events, building on established psycholinguistic methods (e.g., Ellis, 1985–1987). With respect to complex cognitive processes, CODA enhances the widely used research paradigm of using think-aloud protocols and retrospective reports for the identification of (internal) cognitive processes (Ericsson & Simon, 1993). The present approach builds on previous work in this well-established (yet much disputed) tradition, and extends it by suggesting linguistically informed analysis procedures to capture relevant conceptual phenomena reflected in linguistic structure, such as those exemplified above.

CODA as a generic methodology is characterized by essential considerations that lead toward a range of procedures available for data collection and analysis. Depending on the specific aims in a research study, the analyst will need to focus on limited aspects of the linguistic data, since examining unconstrained natural language exhaustively is typically neither feasible nor desirable. This paper will provide the basis for this by guiding researchers through generic linguistic analysis procedures, providing examples for specific analysis perspectives along the way. Following a brief outline of the interdisciplinary background relevant for the CODA methodology, each procedural step will be addressed in turn. Based on a concise presentation

---

[1] Projects where CODA has been explicitly employed as a research method include the Tandem project 'Wayfinding strategies in behavior and language', funded by the Volkswagen Foundation, projects I5-[DiaSpace] and I6-[NavTalk] as part of the CRC (SFB/TR 8) Spatial Cognition Bremen/Freiburg funded by the DFG, and the doctoral project 'Psychopragmatic analysis of problem solving processes', funded by the Zentrale Forschungsförderung (ZF) at the University of Bremen, Germany. Related publications will be referred to throughout this paper. Funding is gratefully acknowledged.

of a range of outcomes, the contribution of language analysis to issues in cognitive science will then be discussed.

## 2. Background

Language has always been one of the core areas in cognitive science, both with respect to its role as a (possibly distinct) cognitive module along with vision, memory, etc. (see Newcombe & Ratliff, 2007, for discussion), and with respect to the relation of language(s) to thought (e.g., Evans & Green, 2006; Langacker, 2000; Talmy 2000, 2007), following Whorf (1941). Concerning the former issue, cognitive scientists (or psycholinguists) are interested in how language is processed in the brain, and how this relates to other (non-linguistic) representations. This includes theories about how languages are learned, how an utterance's meaning can be understood from the acoustic signal and transformed, for example, into a mental image, and how a speaker gets from a non-verbal idea to a linguistic representation.

Concerning the latter issue, cognitive scientists (or cognitive linguists) investigate the features of a language (i.e., the linguistic repertory) with respect to the ways in which it reflects cognitive phenomena. In this area, the identification of systematic differences between languages is central, related to questions about the influence of a language on patterns of thought (in speakers of that language). A major subject of debate is the question of whether (or to what extent) language determines thought (Whorf, 1941), or whether thought is essentially independent of language, and of the language a person speaks. In the latter view, thought patterns determine patterns in language rather than vice versa (Pinker, 1994). Current cognitive linguists appear to converge on a moderate view that allows for dynamic mutual interaction between language and thought (Evans, 2014). Rather than one determining the other, speakers are influenced by the patterns of their language (Boroditsky, 2009), and the patterns in a language develop and are acquired based on its speakers' concepts and usage in embodied everyday activities (Barlow & Kemmer, 2000; Tomasello, 2003). These two aspects of the relation between language and thought are now increasingly seen as complementing rather than contradicting each other.

Both ways in which language plays a role in cognitive science are relevant as starting points for CODA, and there are some shared methodological concerns. Nevertheless the approach presented here is novel in crucial ways, as it departs from established procedures and perspectives.<sup>2</sup> Psycholinguistic

---

[2] This is not to suggest that there are no previous studies that fit to the scope of CODA as presented here. In fact, the idea of developing CODA as a coherent methodology has been inspired by a range of studies cited throughout this paper.

study designs can be fairly similar to CODA, but they aim at optimal control and predictivity rather than freely produced language, since their focus is on cognitive PROCESSING rather than REPRESENTATION. Study designs in the tradition of cognitive linguistics, in contrast, involve examination of the REPERTORY of a language (or languages) with respect to the underlying cognitive representations. Here, a basic tenet is that principles of linguistic structure can serve to reveal principles of cognitive structure, based on the mutual influence of language and thought as just described. The motivation for CODA is that this structural idea carries over to language in use: what we say (and how we say it) is systematically related to, or based on, what (and how) we THINK. This applies not only GENERALLY to what we can do with language or how the linguistic repertory represents the thought repertory within a speech community, but also SPECIFICALLY to what we actually do with language whenever we express our thoughts. In other words, patterns in language USE reflect patterns of CURRENT thought in systematic, though not necessarily direct or unfiltered, ways.

The aim in CODA is to utilize this idea to address research questions in cognitive science, by adopting discourse analytic methods of examining how language is used, and building on previous methods as described above and summarized in Table 1. Unconstrained language, collected in carefully controlled settings, provides a fantastically rich data resource, revealing the ways in which speakers conceptualize crucial aspects of the setting. This includes aspects that the speakers are not necessarily aware of, such as some of the details discussed in the introductory examples. In the following sections, CODA will be introduced procedurally, step by step: this starts from considerations about the scope of using CODA, and further involves data collection techniques, the preparation of data for analysis, drawing insights from content, linguistic feature annotation, concerns of reliability, identification of patterns in the data, as well as triangulation and extensions. While various references to previous studies using CODA will be integrated to support the description, one example in particular (Tenbrink & Seifert, 2011) will be used throughout for illustration. This will be followed by a brief representation of prominent outcomes of CODA-based studies, providing the basis for a critical discussion of the approach.

### 3. CODA procedures

#### 3.1. SCOPE

Cognitive science researchers are interested in a subject that is not directly accessible to observation: processes in the mind and brain, thoughts and thought processes. Although language is an everyday medium used to express thought, there are nevertheless limits to the scope of research that can be

TABLE 1. *Overview of related fields and approaches*

Field / Approach	Relation to CODA
Cognitive science	CODA is a methodology developed to address research questions in cognitive science.
Discourse analysis	CODA uses established discourse analytic methods to address types of research question that have so far rarely been addressed by discourse analysts.
Verbal protocol analysis	CODA enhances the paradigm established by Ericsson and Simon (1993) by analysis of linguistic structure, and by extending the range of text types to be analyzed, taking into account research in cognitive linguistics and communication related aspects.
(Cognitive) linguistic theory	CODA draws on insights from linguistic theory, especially cognitive linguistics, but also functional grammar and other relevant resources, regarding the significance of particular linguistic choices from the overall network of options.
Psycholinguistics	While some psycholinguistic research overlaps with CODA, CODA means eliciting unconstrained language to examine the speakers' linguistic choices, while psycholinguistic studies tend to constrain the range of linguistic options available to participants, typically to address processing aspects.

addressed through language data analysis. As a first step, therefore, it needs to be clarified to what extent language is a suitable medium to convey insights relevant to the research question at hand.

As a starting point, a simple heuristic when considering the scope for CODA is this. Anything that can be meaningfully verbalized by speakers, can be meaningfully analyzed using systematic linguistic methods. As already indicated, the phenomena that a researcher can identify by a close look at linguistic choices may go beyond whatever the speakers themselves verbalize explicitly or would be aware of. Systematic linguistic analysis can thus run deeper than conscious awareness, but it cannot exceed the data resource itself – it cannot address cognitive processes that do not have any reflection in language at all. Consequently, the range of research questions that can be meaningfully addressed via language is mostly limited to non-automatic cognitive processes, excluding memory retrieval, activation and recognition, automated procedures, sudden insights or realizations, and the like. Some of these can be reported after the fact to the extent that they leave a trace in short-term memory (Ericsson & Simon, 1993). Moreover, think-aloud data may to some extent reflect cognitively crucial moments not by explicit formulation but by subtle features such as hesitation markers, pauses, changes in intonation, and the like. Nevertheless, many research issues of interest to cognitive science researchers concern unconscious levels of cognitive processing that will

not find any reflection in language and will therefore need to be addressed in different ways.

While the general scope of potential applications is more extensive, related research traditions fall into two main areas, both of which are combined and extended by CODA. The first tradition concerns *MENTAL REPRESENTATIONS*, namely the linguistic representation of conceptualized information, such as perceptually available or memorized scenes. Typically, this involves description tasks related to an experimentally controlled scenario presupposing no particular cognitive effort (except memory, if the scene to be described is no longer perceptually available). Here a close analysis of linguistic detail is fairly common in order to address the mental representation of perceived information, leading to a broad variety of significant insights (e.g., Berman & Slobin, 1994; Nuyts & Pederson, 1997; Taylor & Tversky, 1996).

The second area concerns the analysis of *COMPLEX COGNITIVE PROCESSES* as identified by verbal protocols produced along with cognitively challenging tasks (Ericsson & Simon, 1993), such as problem solving (Newell & Simon, 1972) or decision-making (Ranyard, Crozier, & Svenson, 1997). The method of having people *THINK ALOUD* during such tasks, or provide a *RETROSPECTIVE REPORT* of what they were thinking, is based on early work by Newell and Simon (1972), and has over the past decades been widely used across many different domains, including spatial cognition (Gugerty & Rodes, 2007), medical areas (Kuipers, Moskowitz, & Kassirer, 1988), design studies (Purcell & Gero, 1998), reading research (Afflerbach & Johnston, 1984), usability (Krahmer & Ummelen, 2004), and many more. Linguistic data of this kind can be seen as an external representation of some aspects of what is going on in the mind. In particular, think-aloud protocols and retrospective reports provide procedural information that complements other data, such as decision outcomes and behavioral performance results.

Conventionally, the focus of verbal protocol analysis lies on the content of verbal data, addressing those aspects (e.g., particular thought processes or strategies) that the speakers are themselves aware of (or 'heed'; Ericsson & Simon, 1993). The content-based inspection of verbal reports, particularly if carried out by experts in the problem domain and set against a substantial theoretical background (Krippendorff, 2004), often leads to well-founded specific hypotheses about the cognitive processes involved (see, for instance, a detailed script analysis in Kuipers et al., 1988).

However, the implications of particular choices of linguistic structure have rarely been taken into account in this line of research. Here, CODA provides a substantial step forward by pointing to the cognitive significance of specific linguistic features, operationalizing and validating content categories, and extending the scope of research to cognitive aspects that are not necessarily

consciously available to the speakers and therefore do not get to be verbalized explicitly. This opens up avenues for addressing research issues that have not yet been investigated using verbal data at all, ideally through triangulation with other kinds of data.

To introduce our running example illustrating CODA procedures, Tenbrink and Seifert (2011) examined how speakers devise route plans for a holiday round trip on Crete. Previous research had highlighted a range of conceptual strategies available for planning complex routes. Hayes-Roth and Hayes-Roth (1979), for example, examined abstract processes reflected in verbal protocols and proposed a cognitive model of planning on this basis. Further research examining scenarios related to the well-known Traveling Salesman Problem (TSP) task was mainly based on behavioral results along with computational modeling, allowing for conclusions about underlying strategies such as clustering mechanisms (e.g., Graham, Joshi, & Pizlo, 2000) or nearest neighbor heuristics (e.g., Best & Simon, 2000). Now, Tenbrink and Seifert (2011) investigated the extent to which results from this research extended to holiday tour planning by examining trajectories, cognitive focus, and conscious strategies for designing a route. Relevant cognitive processes were deemed likely to be verbalizable, related to previous research (e.g., Hayes-Roth & Hayes-Roth, 1979; Tenbrink & Wiener, 2009). Furthermore, Tenbrink and Seifert (2011) were interested in the challenge of conceptually shifting between a two-dimensional map in small-scale space (used for planning) and the three-dimensional real-world situation in large-scale space that travellers are confronted with (see Pick, Heinrichs, Montello, Smith, & Sullivan, 1995, for a verbal protocol study of this challenge). This domain discrepancy is an integral part of any kind of in-advance travel planning. The time and location where planning takes place does not correspond to the time and place of traveling; to ensure a pleasant journey, it is necessary to imagine what kinds of consequence each decision during planning might have for the travel in the real world. While it was unlikely that participants would directly comment on this challenge, the required conceptual shifts should be represented in the language used to describe the problem-solving procedure, in ways yet to be explored. In a nutshell, the research questions for this study were:

- What kinds of problem-solving strategies are used when planning a holiday route with multiple goals, and how do they relate to TSP-related research?
- To what extent do humans refer to a real-world environment when the only available information is a map, and how do they switch between concepts of map-related planning and concepts of traveling in the real world?
- How do these two aspects relate to each other?

## 3.2. DATA COLLECTION TECHNIQUES

Having specified a research question that fits to the scope of CODA as just outlined, the next consideration concerns how to collect data and prepare them for analysis. Since CODA addresses linguistic (and, relatedly, conceptual) patterns by examining HOW some content is expressed or structured (beyond WHAT is said), speakers (i.e., the participants of empirical studies) need to be allowed to make their own linguistic choices rather than choosing from a restricted set of options. The specific elicitation methods used in a study need to be chosen carefully based on the research question at hand, keeping in mind that each particular text type implies particular patterns of linguistic choices (e.g., Biber, 1989) that may not be related to the given task as such.

In the area of mental representations, the elicited verbal description should optimally reflect the speakers' conceptualization of a perceived scene or event. For example, participants could be shown a picture and asked a question about it that triggers a description (e.g., Carlson & Logan, 2001; Henderson & Ferreira, 2004; Holsanova, 2008). Notably, the precise formulation of the question and other discourse factors systematically affect the participants' description. As Vorweg and Tenbrink (2007) showed, a question asking about the LOCATION of an element in a picture ("Where is the object?") triggers far more detailed descriptions than a question about the IDENTITY of an element ("Which one is the object?"), although both questions can generally be answered in similar ways, as exemplified by "it is (the object) to the left of the square". Responses to 'Where' questions contained more projective terms and modifiers, as in "slightly to the top and left of the square", whereas responses to 'Which' questions tended to be short and simple, as in "the one to the left" or simply "the circle".

In the area of problem-solving processes, Ericsson and Simon's (1993) framework provides a good basis for identifying the cognitive significance of certain text types. Most prominently, information verbalized during the task (THINK-ALOUD protocols) and RETROSPECTIVE REPORTS are supposed to reflect cognitive processes within short-term memory fairly directly, and can therefore be recommended as preferred elicitation methods. Ericsson and Simon (1993) provide elaborate procedural advice towards optimal elicitation for both methods. The main idea is to encourage participants to speak out loud what they are (or – in the case of retrospective reports – WERE) thinking, rather than guiding them towards particular trains of thought. The 'Appendix' provides an example instruction that can be used to train participants to think aloud, which is necessary if only to clarify what is expected from them.

It has been observed that, under certain circumstances, the requirement to verbalize may promote a better understanding of the task itself

(Krahmer & Ummelen, 2004) – or it may lead to an impairment (Schooler, Ohlsson, & Brooks, 1993). Based on a range of studies showing effects in either direction (or none), it now seems clear that it cannot be generally predicted whether the requirement to think aloud will affect task performance in a particular research setting. Therefore, studies using think-aloud protocols typically involve another (control) group of participants who are asked to perform the same task without thinking aloud, allowing for a comparison of behavioral results.

Additionally, relying on think-aloud data alone may often not be sufficient since verbalizations during the task may be incomplete in various respects (Ericsson & Simon, 1993). Other types of verbalizations have different effects. Therefore, it may be useful to combine several methods of data collection, both with respect to other types of verbal data and with respect to triangulation (addressed separately below). In the following, I will briefly address some other widely used types of language data elicitation, which may be suitable for different purposes.

### 3.2.1. *Interview questions*

One very direct way of eliciting responses of interest to the researcher is by asking people direct questions about their experience concerning a task just performed. Such questions are often formulated in such a way as to differentiate between alternating theories (e.g., Schelhorn, Griego, & Schmid, 2007, with respect to analogical reasoning strategies) and can therefore be quite specific and conceptually biased. Then the formulation of interview questions may not necessarily map onto the participants' personal experience of the task. In particular, Ericsson and Simon (1993) point out that questions posed by the experimenter, if not formulated in a very general way, lead to filtering processes and may address aspects that the participants never actually attended to by themselves during the problem-solving process (such as reasons and motivations). While researchers should be aware of these effects, they may be used to advantage, e.g., by complementing other (less biased) kinds of response (e.g., Gralla, Tenbrink, Siebers, & Schmid, 2012).

With respect to problem-solving tasks, thought processes triggered by interview questions can lead to the mention of strategies that could have been used but were not. Due to conscious reflection, participants may realize that better performance on the current task could have been achieved. Such recognition of further possible strategies would in most cases also be reflected linguistically, for example by discourse markers that signpost the new insights gained through the interview. Again, this highlights the need for a close examination of the language used, beyond extracting the types of strategies mentioned by participants.

### 3.2.2. *Scenario variation*

Another elicitation method is to suggest different discourse tasks or scenarios to the participants. This involves eliciting verbal representations not only for the purpose of revealing thought processes, but primarily for a different purpose in which these thought processes are again put to use, this time in order to create a linguistic product. For example, in Tenbrink and Wiener (2009), as well as Gralla et al. (2012), participants were asked first to provide a retrospective report of how they solved the problem given to them, and then to write an instruction 'for a friend' (a new discourse task), leading to new conceptual perspectives on the task at hand. In a route-planning task, Hölscher, Tenbrink, and Wiener (2011) had people, in one condition, describe their future route for themselves, and in another condition, for a stranger unfamiliar with the environment, highlighting systematic quantitative differences (level of detail) but striking qualitative correspondences (same types of information given). Relatedly, Daniel and Denis (2004) asked participants to give route descriptions either normally or in a specifically concise way, thus identifying systematic features of condensed route descriptions.

Similarly, revealing insights can be collected through *DIALOGUES*, for instance involving participants with equal or different levels of knowledge. Involving real or imagined addressees shows how experiences in mentally representing a scene or solving a problem may be shared communicatively, or how cognitive processes can be conveyed from an expert (in solving a complex task) to a novice. Clark and Krych (2004) present a relevant analysis of dialogues concerned with a joint problem-solving task (building a LEGO model), showing how experts adjust their instructions according to their partners' reactions. In route directions, participants use a verbal representation to enable another person to find their way (e.g., Denis, 1997), which opens up further possibilities for eliciting language under consideration of different perspectives. Apart from the text type itself, the precise nature of the (perceived) discourse goal (i.e., why language is produced) plays a decisive role, which influences the trains of thought that are triggered by the way the current linguistic aims are understood. Clearly, it matters to the participant whether they perceive a description to be for the experimenter only (which would inevitably be the case with imagined addressees), or for successful communication to be actually required for the given discourse task (Schober & Brennan, 2003). Both cases, however, involve a higher involvement of communication related aspects than think-aloud protocols, which (ideally) are not primarily directed at anybody at all.

Like interview questions, instructions for other people (imagined or present) as well as dialogues may trigger intermediate processes of verbalization,

such as explanations. They need to be understood as going beyond a direct representation of thought, and can provide insights about participants' meta-conceptualization and rationalization of their choices.

In terms of our example study, Tenbrink and Seifert (2011) elicited written reports of the problem-solving procedure. They asked participants to write down what they did, step by step, when designing a holiday trip on Crete, what their thoughts were, and what was important while making decisions. This kind of verbal report is not a direct representation of thought as recommended for retrospective reports by Ericsson and Simon (1993), yet it served the study purposes by providing insight into the participants' conscious thoughts and meta-cognition in this task. Moreover, the written mode guaranteed sufficiently rich verbal data to analyze the language used with respect to conceptual shifts between planning and traveling domains (map as opposed to real world).

### 3.3. DATA PREPARATION TECHNIQUES

Following data collection, the next procedural step is to prepare the verbal data for analysis by transferring them into manageable units. Handwritten language needs to be transferred to an electronic format; spoken language needs to be transcribed. If transcription is involved, the required level of detail and types of feature to be represented in the protocol need to be considered carefully. Since CODA typically addresses semantic or conceptual levels of linguistic structure (see Section 3.4), an orthographically correct transcription of what is meant may be more useful than an exact representation of pronunciation patterns. However, it is typically useful to transcribe markers of hesitation (such as *hm*, *uh*, *uh-uh*) and the like in systematic ways. Such sounds are not produced at random; they usually carry a meaning relevant for communicative and/or current cognitive processes. Analyzing them systematically can be enlightening for research areas involving uncertainty (Lindsey, Greene, Parker, & Sassi, 1995; Tenbrink, Bergmann, & Konieczny, 2011) or confusion, cognitive effort, dialogic negotiation, and other issues (Brennan & Williams, 1995). Regardless of whether or not a systematic analysis of hesitation markers is carried out, including this information in the transcript supports the interpretation of utterance meaning.

To represent discourse functions of intonation in a feasible and well-established (though coarse) manner, punctuation markers can be used in the conventional way, i.e., question marks for (semantically identifiable) questions, exclamation marks for the (unusual) case of an exclamation, commas for a continuing intonation contour even if a sentence is grammatically complete, and a full stop to signal ostensive completion (e.g., as signaled by a falling

intonation contour). Non-linguistic events can be noted in brackets, such as (*laughter*) or (*noise*). The convenient transcription software f4<sup>3</sup> uses time stamps; this allows for capturing the temporal development as well as extended pauses. Short and mid-utterance pauses can be represented in brackets using numbers for seconds. More intricate non-verbal contributions such as gestures and facial expressions should be considered with respect to their relevance to the research goals; transcribing these can be extremely time-consuming, especially if no established conventions for the particular distinctions needed are known (see Brösamle, 2013, for an extended specialized gesture transcription project). Furthermore, actions may be crucial for the interpretation of language. Generally, all relevant information needs to be included on a suitable level of detail. Here, the focus will remain on language.

Once the language data are available in electronic format, they will typically need to be segmented into smaller units serving as the basis for analysis. The length and definition of a unit depend on the research goals (see Krippendorff, 2004, for elaborate discussion of unitization). Any kind of quantitative analysis must build on carefully defined (operationalized) units. However, this goal may not always be easy to achieve for smaller units of analysis. In some cases it may be sufficient to establish smaller units simply for practical purposes, e.g., as a basis for line-by-line annotation (see Section 3.4). Researchers may decide to relate quantitative analysis results to the overall number of words or other clearly delineated larger units, while retaining smaller units simply for purposes of managing the annotation process.

A notion that may be a useful start for line-by-line analysis is that of a ‘possible sentence’ (Selting, 2000). Selting specifies this notion in terms of **TURN-CONSTRUCTIONAL UNITS (TCUs)** as follows:

The TCU is defined as the smallest interactionally relevant complete linguistic unit, in a given context, that is constructed with syntactic and prosodic resources within their semantic, pragmatic, activity-type-specific, and sequential conversational context. (Selting, 2000, p. 477).

While Selting developed her notion of TCUs for spoken language, a similar idea can also be used for written language in order to obtain manageable units of a similar size; these may be shorter than the sentences suggested by the writer’s use of punctuation. Moreover, in dialogue, **TURN CHANGES** between speakers are clear cases of new units, and **SYNTACTICALLY COMPLETE SENTENCES** are also units. **SEMANTIC/PRAGMATIC COMPLETION** is reached, for instance, when a speaker turns from a description of a specific item or spatial array to the next. For some

---

[3] Online: <<http://www.audiotranskription.de/english>>.

purposes the notion of *conversational game* established in the dialogue modeling literature may be useful (Carletta, Isard, Isard, Kowtko, Doherty-Sneddon, & Anderson, 1997). Further useful ideas about segmentation and the definition of discourse units can be found, for instance, in Allen (2000), Degand and Simon (2009), Denis (1997), and Krippendorff (2004). However, for many purposes it may not be necessary to identify a specific operationalization of unit definitions, especially if no quantitative analysis directly relies on unit counts. Then, segments can be intuitively defined as convenient for the analyst.

Tenbrink and Seifert (2011) first rendered the collected handwritten reports into digital format. Unitization was done on the basis of informational chunks, similar to the notion of TCU as described above (Selting, 2000, i.e., “the smallest interactionally relevant complete linguistic unit” in the given context), as exemplified here:

1. Zentralen bzw. großen Ort mit Flughafen gesucht (Sitia)  
Searched for a central or large town with an airport (Sitia)
2. Dann habe ich überlegt wie weit man in 2 Wochen so kommen könnte.  
Then I considered how far it is possible to travel in two weeks
3. Ziel war es möglichst viel zu sehen  
The aim was to see as much as possible
4. und weit in den Westen zu kommen.  
and to get far into the west.
5. Große Städte als Übernachtungsplätze:  
Large cities for overnight stays:
6. viel Unternehmungsmöglichkeiten [sic!] für Touristen  
many activity options for tourists

This unitization made it possible to analyze chunks of text in a straightforward line-by-line analysis.

#### 3.4. CONTENT ANALYSIS

Following data preparation, the first step of any analysis of natural language data is to gain a clear grasp of the content of the data, i.e., the speakers' meaning in producing the verbalizations. An intuitive understanding of the range of content produced by the speakers should precede any closer analysis as described next. Based on the content and guided by the research question that motivates the study at hand, the analyst will need to make decisions about which aspects to pursue further and capture systematically.

Content analysis, as described by Krippendorff (2004) and Ericsson and Simon (1993), can involve extremely complex analysis procedures (see Crampton, 1992, for an insightful example). In particular, Krippendorff (2004)

describes content analysis as a research technique suitable for making reliable inferences from texts; this represents a notion that considerably exceeds simple, superficial, intuitive text comprehension. In problem-solving studies, it is often possible to identify a range of conceptual strategies, representations, and processes that are directly described by participants (cf. Ericsson & Simon, 1993; for CODA-related studies see, e.g., Gralla et al., 2012; Hölscher et al., 2011; Tenbrink & Wiener, 2009). Moreover, content analysis provides a first basis for categorizing the data with respect to each segment's relation to the discourse task (related to the research question at hand). For instance, in Tenbrink, Coventry, and Andonova (2011), utterances were categorized as to whether they described an object's location or orientation, both, or neither. These categories were identified because they emerged as prominent types of content produced by the speakers, with clear effects on the targets of the research design.

In Tenbrink and Seifert (2011), content analysis served to inspire the more fine-grained coding procedures (as described in Section 3.5). While examining how participants described the procedures of planning a holiday tour and how they shifted between the conceptual domains of planning and traveling, it became clear that the participants sometimes explicitly commented on the former issue, while the latter remained implicit in language. To capture explicit comments, the authors extracted any mention of conscious strategies describing the holiday tour design procedure. Examples are: "After that, I also looked for a 'peaceful' place with a range of attractions in the vicinity", which exemplifies the importance of spatial vicinity for travel planning; and "It was also important to avoid traveling the same route twice, but rather, traveling some sort of circle (ellipse)", which exemplifies the concept of a suitable overall trajectory as well as the avoidance of repetition. As a next step, linguistic indicators for each of the identified strategies (e.g., *region*, *area*, *side*, and *mainland* indicate a conceptual segmentation of the environment into regions) were determined in order to operationalize the strategy allocation process (see Tenbrink & Seifert, 2011, for details).

Concerning domain shifts, the general impression emerged that participants were mentally at two places at once rather than showing awareness of a conceptual shifting process. A closer inspection of the language the participants used led to the identification of linguistic markers indicating each of the domains involved (the current conceptual domain of planning alongside that of traveling on a remote island) as well as the (implicit) shifts between them. The details of this part of the analysis will be described in Section 3.5.

### 3.5. ANALYSIS OF LINGUISTIC FEATURES

Building on the identified content categories, the next step in the analysis is to identify features of the linguistic representations that can be interpreted as

reflecting characteristic conceptual phenomena. As illustrated above, some aspects of language reflect cognitive aspects that go beyond conscious reflection by individual speakers, and that are not necessarily directly observable in linguistic content. Speakers are typically unaware of the cognitive structures that are reflected in particular ways of framing a representation linguistically. Furthermore, they are not consciously aware of the NETWORK OF OPTIONS (or ‘social semiotic system’; Halliday & Matthiessen, 1999) that allows for a range of linguistic choices beside their own. For instance, a sentence like *The car is next to the tree* will be intuitively produced without considering alternative options like *The tree is next to the car*, *The oak partially covers the Bentley*, or other linguistic representations of the same scene. Nevertheless these choices are meaningful. From a cognitive point of view (Talmy, 2000), *The car is next to the tree* is more standard than *The tree is next to the car*, since the movable object (car) is referred to as a locatum in relation to the fixed object (tree) as relatum. Besides cognitive principles, discourse-related factors (such as the current topic of the conversation) may lead speakers to intuitively choose other options. Depending on the context of its production, a choice like *The oak partially covers the Bentley* may reflect the speaker’s way of perceiving a pictorial configuration as well as their attention to details (*oak* rather than *tree*, *Bentley* rather than *car*). Along these lines, different ways of referring to the same situation reveal the speaker’s conceptual perspective (Schober, 1998), without the speaker necessarily being aware of this effect.

To capture these issues, it is useful to first examine the data qualitatively so as to identify linguistic features that are relevant to the research issues at hand. This means a close look at the ways in which central aspects relating to cognitive processes are expressed in language. For instance, it may be interesting to examine whether decision points in route descriptions are linguistically represented as given, and backgrounded, or rather highlighted as new elements. An examination of the language used to refer to these locations will then lead to the identification of the relevant linguistic repertory, such as definite and indefinite articles, modifiers, syntactic position, and the like.

The next step is then to annotate the data on a line-by-line basis so as to capture crucial qualitative insights systematically. Apart from gaining quantitative insights by counting numbers of occurrences of particular phenomena, systematic patterns can then be identified based on the features’ distribution throughout the data (see Section 3.6).

This approach to linguistic data analysis is fundamentally discourse-analytic. Discourse analysis generally means analyzing texts with respect to their linguistic (and contextual) features, adopting a specific analysis perspective that is relevant with respect to a particular motivation – bearing in mind that linguistic analysis can almost never be regarded as exhaustive.

Crucially, in contrast to psycholinguistic experimentation, which typically relies on precise predictions and controlled settings, the identification and detailed (qualitative) description of relevant linguistic phenomena and their interpretation relative to the research question is seen as primary. Quantitative data then serve to highlight the relative role of the detected phenomena within the text, and further support can be gained through inferential statistics (e.g., comparing results for different conditions; see Section 3.6).

In other areas of discourse analysis, research motivations include identifying distinctive features of text types (de Beaugrande, 1980), e.g., for purposes of data mining (Leidner & Schilder, 2010), specifying dialogue structure (e.g., to inform automated dialogue systems, as in Shi, Jian, & Rachuy, 2011), and uncovering gender bias or a hidden political agenda (as in *Critical Discourse Analysis*, e.g., van Dijk, 1993). In CODA, the perspective adopted is to identify linguistic features that are potentially indicative of cognitive processes and representations.

It is crucial at this central procedural step to provide clear and crisp definitions and operationalizations for annotation criteria, typically aiming for mutually exclusive variables within a category (see Carletta et al., 1997, for a particularly useful discussion of these issues for the area of dialogue annotation). In the following, I outline some examples for selective discourse analysis perspectives relevant for research in cognitive science, motivating systematic annotations of specific linguistic features in the collected data.

First, the way in which texts (of any type) are linearly structured, and the way in which the information is presented, can be expected to relate systematically to the way the underlying cognitive processes are structured. This concerns both the text as a whole, revealing, for instance, temporal and causal relationships developing gradually, and smaller portions of the text, for example information packaging within single clauses (Halliday, 1994). For instance, with respect to whole texts, Tenbrink and Wiener (2009), as well as Gralla et al. (2012), exploited the temporal structure of participants' problem-solving reports for the purpose of proposing a generalized procedure for the type of problem at hand. These accounts served to illustrate how different cognitive strategies, which in earlier literature were treated as contrasting problem-solving approaches, were integrated over time by the problem-solvers in the studies. With respect to smaller units of text, Tenbrink and Ragni (2012) highlighted regular principles within participants' ways of describing abstract configurations, both on the level of whole configuration descriptions and with respect to individual analysis units that described the relation of two objects to each other. This analysis revealed systematic visual attention patterns with respect to small-scale space. Rather than conceptualizing a configuration as a whole and describing its details in random order, people mostly followed conventional patterns of reading (left to right),

and furthermore tended to base each individual object location description on the previous one (e.g., *A is at the top left, B is next to A, and C below B*).

Second, for many purposes, a close look at the speakers' lexical choices in references to objects and notions can be revealing (see also Krippendorff, 2004). Since there may be a range of different options, the particular reference type chosen by a speaker highlights the role of a particular semantic or conceptual field within the current verbalization task. In many cases, specific lexical items that have been identified by cognitive linguists (e.g., Talmy, 2000) as reflecting cognitive structure may be indicative of underlying conceptual patterns. For example, explorations of the chosen level of detail (Daniel & Denis, 2004; Vorwerg & Tenbrink, 2007), hierarchical description levels (Plumert, Carswell, de Vet, & Ihrig, 1995), and the underlying conceptual perspective (Tenbrink, Coventry, & Andonova, 2011; Tenbrink, Ross, Thomas, Dethlefs, & Andonova, 2010; Tversky, 1999) reveal the flexibility of these concepts relative to changes in the task scenario. For example, Plumert et al. (1995) showed how the discourse task as well as the spatial configuration of landmarks affected the order and hierarchical structuring of spatial descriptions. Tenbrink et al. (2010) found that speakers interacting with other humans frequently shift between perspectives in route dialogue, whereas speakers communicating with an automatic dialogue system refrain from doing so and mostly stick to the (arguably) simpler perspective choice.

Moreover, it can be useful to trace the development of reference types over time, particularly for complex cognitive processes. If reference types change during a problem-solving task, this "can be interpreted as the trace of changes in the functional organization of the subject's representation" (Caron, 1996, pp. 24f.). A detailed analysis of the semantics of the lexical choices in the relevant action context may highlight the significance of this kind of conceptual change. For example, in a comparison of think-aloud protocols and instructional discourse in a dollhouse assembly setting, Gralla (2014) shows that reference choice is influenced by prior knowledge about the function of referent objects. Speakers who had been shown a picture of the fully assembled dollhouse often used pronouns in initial reference to a part in focus, reflecting its integration into an existing mental representation. In contrast, participants without prior knowledge initially tended to use definite noun phrases containing domain-unspecific nouns. Moreover, reference choice was also influenced by gradual change in the comprehension of the situation. Subsequent references reflected mental re-conceptualizations via the assignment of specific functions to objects. These principles, which systematically affected the distribution of pronoun use as well as lexical choices, were mediated by the communicative purposes in an instruction context.

Third, it is often worthwhile to examine the use of explicit discourse markers, which may serve multiple purposes, and have been intensely

researched both for English and for German (Fischer, 2006; Grosz & Sidner, 1986; Schiffrin, 1987). According to Caron (1996), the use of discourse markers in think-aloud protocols reveals how the participant construes the concepts and relations involved, without serving any specific communicative purpose that might influence this construal. For example, connectives (such as *before*, *because*, *while*) explicitly structure the represented contents. Furthermore, certain markers that are particularly prominent in spoken language may reflect hierarchical thought processes (see also Bégoïn-Augereau & Caron-Pargue, 2003); for instance, occurrences of *Okay*, *now ...* may signal the completion of a subprocess together with the start of a new one. Caron (1996) specifically proposes that, if modal expressions such as *can*, *must*, *have to* occur together with interjections such as *oh*, *well*, this

can be interpreted as traces of operations by which the subject does not work anymore on the current representation, but ‘withdraws’ from it [...] in order to have access to another representation. It may correspond either to the planning of a new course of action or to the access to knowledge stored in long-term memory. (Caron, 1996. pp. 25f.)

Other verbal cues, such as pauses, lapses, and self-repairs, may be indicative of other types of cognitive process, depending on the task situation (Lindsey et al., 1995; Tenbrink, Bergmann, & Konieczny, 2011).

Generally speaking, a detailed analysis of linguistic features is the most central aspect of CODA. There is no theoretical limit to the types of linguistic feature that can be or should be attended to in this part of the analysis. After all, language relates to cognition in many different ways, as shown by linguistic theory and previous research. Crucially, linguistic features need to be identified that are relevant for the research question at hand. The analyst will need to take account of practical limitations; while many aspects may be interesting to examine and discuss in depth, it will typically only be feasible to pick a few of them and focus on these. Identifying the most relevant features in a linguistic dataset may take time and effort, but will prove worthwhile if the analysis is then carried out systematically.

In terms of our running example, Tenbrink and Seifert (2011) identified switches between different conceptual domains by classifying lexical choices in the collected retrospective reports. In particular, they examined the linguistic choices in each analysis unit with respect to the following features:

- identification of an underlying (explicit or implicit) AGENT of a described action that could be identified unambiguously as a PLANNER or TRAVELER (e.g., due to the nature of the instruction in this task, *I* was typically the planner, whereas *they* referred to traveling);

- nouns, verbs, and adjectives/adverbs that could be identified as indicators of the **PLANNING** or the **TRAVELING** activity (where *map*, *decide*, and *important* were typical indicators of planning, while *relaxation*, *experience*, and *beautiful* referred to traveling);
- **TEMPORAL MARKERS** that clearly belonged to the planning or the traveling domain (identifiable by the temporal scope involved, as illustrated by references to *days* and *hours*).

In terms of a line-by-line analysis of the units cited above, this looks as shown in Figure 1. Importantly, each annotation category was carefully defined in an annotation documentation file so as to avoid confusions and inconsistencies in unclear cases. For instance, places and activities that were directly represented in the map (such as the ‘large town’ Sitia) had a corresponding element both in the planning and in the traveling domain, and therefore could not be regarded as indicating either of the domains, explaining the zeros in line 1 (Figure 1). Similarly, activities (*Unternehmungsmöglichkeiten*) were indicated as symbols in the map. Consequently, the only nouns that were counted as indicators of a specific conceptual domain were *Übernachtungsplätze* ‘overnight places’ and *Touristen* ‘tourists’.

As in this example, the linguistic analysis in Tenbrink and Seifert (2011) started from a qualitative examination of indicators of a particular phenomenon (a cognitive domain or a conceptual strategy), followed by the identification (annotation) of their occurrences in the dataset using number counts in the data table. While this is a frequent procedure in CODA, data tables can also represent and support more complex analysis procedures. Consider the following example, taken from Vorwerg and Tenbrink (2007), where level of detail in the description of an element in a picture was addressed by a close examination of linguistic features, as represented in Figure 2. Here, the first two columns provide the number code for the condition and the picture that is described by the utterance represented in the third column. The annotation starts in the fourth column, where mention of direction terms is categorized as follows: ‘1’ for one direction term (e.g., *top* for picture 1), ‘2’ for more than one direction term (e.g., *northwest* and *lower* for picture 2), and ‘0’ for descriptions containing no direction term at all (as for picture 7). The fifth column asks whether the direction term (if any) is modified by a **PRECISIFIER** such as *most* (for pictures 8 and 10); other examples would be *slightly (left)* and *directly (above)*. If the direction term remains unmodified, the code is ‘1’; if it is modified, the code is ‘2’; and if there is no direction term the code is ‘0’. Note that this annotation creates nominal rather than ordinal or cardinal categories. Alternatively, the number of direction terms or precisifiers could have been counted. The last column, however, differs in this respect since it asks for a **TYPE** of relatum

Informational unit	Agent: Planner	Agent: Traveler	Planning nouns	Traveling nouns
Zentralen bzw. großen Ort mit Flughafen gesucht (Sitia) [Searched for a central or large town with an airport (Sitia)]	1	0	0	0
Dann habe ich überlegt wie weit man in 2 Wochen so kommen könnte. [Then I considered how far it is possible to travel in two weeks]	1	0	0	0
Ziel war es möglichst viel zu sehen [The aim was to see as much as possible]	0	1	0	0
und weit in den Westen zu kommen. [and to get far into the west.]	0	1	0	0
Große Städte als Übernachtungsplätze: [Large cities for overnight stays:]	0	0	0	1
viel Unternehmungsmöglichkeiten für Touristen [many activity options for tourists]	0	0	0	1

Fig. 1. Snapshot of the line-by-line annotation of conceptual domains in Tenbrink and Seifert (2011).

(rather than existence or frequency). Here *NO RELATUM* is coded as '0' and *THE BOX* as *relatum* as '1'; alternatives found in the data were *THE SPEAKER* (e.g., *in front of me*) or *THE PICTURE* itself (e.g., *leftmost in the picture*) as *relata*, and these were assigned different number codes. Instead of numbers, it is equally possible to use lexical categories (e.g., *box*, *speaker*, etc.) in the annotation columns.

### 3.6. RELIABILITY

Following line-by-line annotation, the annotations need to be checked for reliability, as outlined, for example, by Krippendorff (2004). This is an established way of assessing to what extent the definitions used for analysis were operationalized sufficiently for use by different annotators. The aim is to rely not only on experts to use these definitions, who typically draw on additional background knowledge that is hard to determine, but also other coders who should not need extensive additional training. Where applicable and feasible, inter-coder reliability should be tested statistically, for example using Krippendorff's alpha (using the tool provided by Hayes & Krippendorff, 2007, for the SPSS/PASW statistics software package). As shown by Krippendorff (2004), the widely used measure Cohen's kappa (Cohen, 1960) has the flaw of allowing for systematic disagreement, which may then go unnoticed.

Tenbrink and Seifert (2011) let independent coders annotate a representative subset of the data in parallel, and obtained favorable Krippendorff's alpha values for each annotation category.

### 3.7. IDENTIFICATION OF PATTERNS

Having accomplished a detailed and reliable annotation of the linguistic features of individual participant data, the results need to be systematically

condition	picture		no. of direction		type of
	no.	response	terms	precisifiers?	relatum
2	1	top element	1	1	0
2	2	circle northwest of lower bar	2	1	0
2	3	circle to the left of the back of the box	2	1	1
2	4	circle behind the box	1	1	1
2	5	circle northwest of the box	2	1	1
2	6	northwest circle (or) upper left circle	2	1	0
2	7	circle	0	0	0
2	8	westmost circle (or) leftmost circle	2	2	0
2	9	circle northwest of the box	2	1	1
2	10	westmost circle (or) leftmost circle	2	2	0

Fig. 2. Snapshot of the line-by-line annotation of level of detail in Vorwerg and Tenbrink (2007).

related to the features of the setting in which the language was produced. This is most conveniently accomplished by transferring counts of annotation results per participant into a spreadsheet. Patterns to be identified may concern the features of a verbally represented scene, or the problem-solving process, different conditions, scenes, or situations, subgroups of participants or individuals, or different text types. For instance, Daniel and Denis (2004) identified the features of route descriptions that were relevant for conciseness in all collected data, and then compared the results between conditions (which differed in the extent to which conciseness was explicitly asked for). Where other kinds of data, such as behavioral results, are available, the analysis results furthermore need to be related to these (see Section 3.8). Establishing such inter-relationships between types of evidence is not only useful in terms of validation (Krippendorff, 2004) but also in terms of accounting for the significance and impact of any kind of change in the situation, its conceptualization, and its representation in language.

In Tenbrink and Seifert (2011), the line-by-line annotation of conceptual domains based on linguistic indicators was used to identify units that referred ONLY to the traveling or planning domain, both, or neither. A close examination of these led to the identification of linguistic markers of conceptual shifts (such as *let* in *I wanted to let the couple travel once round the western part of the island*). Number counts for each category per participant were transferred to the spreadsheet, and then related to various categories relevant for the study, namely different modes of travel for different participants as well as different conceptual planning strategies. Together, these analyses revealed how participants used spatial strategies to design a travel plan relevant for a conceptual domain that is distinct from the currently perceived scene, taking into account scenario features such as the mode of travel involved. There were striking parallels between this naturalistic and conceptually complex holiday planning task and previous results for

simpler and more abstract spatial planning problems. Furthermore, the conceptual domain shifts highlighted by the linguistic analysis resonates with theories on conceptual flexibility, such as that reflected by the adoption of various perspectives.

Depending on the sample size, existence of different conditions, distribution of results, and the like, the analysis may now be supported by statistical procedures. Since the type of data and feature distribution will vary from case to case, no recommendations for specific statistical tests can be made here. Previous statistical analyses related to CODA range from simple *t*-tests, chi-squares, and ANOVAs to more complex mixed-effects logistic regression models.

Obviously, statistical validation of observed patterns is highly desirable and provides substantial support for the linguistic analysis results whenever it can be achieved. It should be recognized, however, that this is notoriously hard to obtain based on unconstrained language data, and statistical significance is not the only valid evidence of cognitive phenomena. In various areas of cognitive science (e.g., artificial intelligence, human–computer interaction, cognitive modeling, and others), it is actually common to examine individual case studies, to obtain proof of concept, and the like. Showing that phenomena exist (maybe systematically under distinct circumstances) can be a decisive step forward in the understanding of the human mind. Case studies and the identification of qualitative patterns can therefore be regarded as inspiring explorative insights, leading towards more controlled study designs that can shed further light on the observed phenomena.

In terms of our running example, Tenbrink and Seifert (2011) focused on descriptive statistics (relative frequencies). Much of the analysis was qualitative, showing the repertory of conceptual strategies as well as domain shifts as represented in language across various scenario types. The quantification provided an assessment of the relative role of these processes within the overall procedures and verbalizations.

### 3.8. TRIANGULATION AND EXTENSIONS

As observed by many researchers following the tradition of verbal protocol analysis (based on Ericsson & Simon, 1993), language may in many ways be insufficient for gaining access to cognitive processes and representations to the extent desirable for a research purpose. A systematic linguistic analysis (as just described) partially remedies this by a closer examination of linguistic choices than available through content analysis, building on established insights concerning their significance. Nevertheless, it is highly beneficial to collect other types of evidence that can complement the insights gained from language, as well as to relate insights gained from verbal protocol analysis to established (or newly developed) theories and models.

To link the results of linguistic analysis with other findings, cognitive science provides a wide range of methodologies and opportunities for triangulation. For example, measures such as memory or behavioral performance data, decision outcomes, reaction times, and eye-tracking data can provide (further) insights on cognitive activities, some of which remain below the threshold of participants' awareness. Eye-tracking behavior, for instance, primarily reveals implicit patterns of attention of which participants are rarely aware (Findlay, 2004; Thomas & Lleras, 2007). Neuroimaging (e.g., fMRI; see Cabeza & Klingstone, 2001, for an overview) provides insights on the activation of particular parts of the brain, for example during the conceptualization of objects, which triggers activation of certain lexemes (Boutonnet, Athanasopoulos, & Thierry, 2012).

Synthetic approaches such as cognitive modeling (e.g., ACT-R; Anderson, 2007; Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004; Anderson & Lebiere, 1998) suggest mechanisms and procedures across all cognitive levels and modules, including those that can only be hypothesized at the current state of research. In particular, cognitive models represent theories to explain the observable data consistently by concrete assumptions about the underlying non-observable processes, covering the full range of cognitive processes from conscious human decisions and strategies through to neural activities. The result of a computationally implemented cognitive model can then be compared with observable insights from various sources, including fMRI (Fincham, Carter, van Veen, Stenger, & Anderson, 2002; Ragni, Fangmeier, & Brüssow, 2010) as well as language data, feeding back into improved models. For example, Gugerty and Rodes (2007) provide an ACT-R model of the strategies and cognitive processes involved in cardinal direction judgments, based on participants' verbalized strategies and enhanced by further findings about human direction conceptualizations and other processes that were not directly reported by participants.

Another extension is to feed the results into practical applications. The overarching goal motivating the study reported in Tenbrink and Seifert (2011), for instance, was the development of spatial assistance software for planning holiday routes (Seifert, 2008). Other studies have focused on the development of intuitive human–robot or human–computer interaction (e.g., Moratz & Tenbrink, 2006; Tenbrink et al., 2010).

### 3.9. SUMMARY

Table 2 provides an overview of the steps involved in CODA along with some key questions to be considered by the researcher, and key aspects involved in a step. Each step can be treated more or less elaborately in a specific research process.

TABLE 2. *Overview of steps in the CODA procedure*

CODA step	Researcher's considerations	Key parameters in CODA
Scope	Clarify if the current research question can be addressed by verbal data analysis. Which aspects of interest to the researcher may be revealed through language?	CODA can provide insights about mental representations and complex cognitive processes.
Data collection	Which type of language is best suited for the research question and experimental design at hand? How should the instruction to participants be formulated?	CODA can be applied to verbal descriptions, think-aloud protocols, retrospective reports, interviews, procedural instructions, dialogues, etc. in spoken and written modes.
Preparation of data for analysis	Which format and level of detail is most suitable for current purposes? What kinds of information are important for analysis? What kinds of units are adequate, considering the content as well as analysis procedures?	Using a suitable electronic format, transcripts should at least represent the words used by the participant, and may involve more details such as pronunciation, hesitations, pauses, etc., depending on the analysis targeted in this research. Non-verbal information such as gestures and actions need to be included if relevant, choosing a suitable level of detail. Units need to be small enough to handle and annotate systematically, and may be based on content or syntax (e.g., sentences).
Content analysis	What are the main contents conveyed in language? Do they fall into distinct categories? Do they lead to intuitions or insights that should be pursued systematically? Are any relevant concepts and strategies conveyed explicitly?	Content aspects can be used as anecdotal evidence, as well as for a systematic analysis of what thoughts are conveyed under what circumstances. A good grasp of the content can lead to the identification of crucial linguistic features to be analyzed.
Linguistic feature annotation	What are the main linguistic features that are relevant for, or revealing of, the cognitive processes addressed in this study? How do speakers express the most crucial content – what are their linguistic choices, and which cognitive structures may they reflect?	This analysis builds on insights (or intuitions) about the significance of specific linguistic features. It requires clear definitions of linguistic feature annotation categories and criteria (ideally building on established linguistic insights and previous literature), and systematic annotation of all relevant language according to these definitions.
Reliability	Check if different coders annotate the data in the same way, given a good understanding of the annotation definitions and as much context as needed.	At least two independent coders should annotate a representative subset of data in parallel. Krippendorff's alpha is a good measure to assess inter-coder reliability.

TABLE 2. (*Cont.*)

CODA step	Researcher's considerations	Key parameters in CODA
Identification of patterns	How do the annotated linguistic features relate to features of the scenario or the individual participants, conditions, procedural steps, etc.? How do actions come into play?	Spreadsheets provide a useful summary representation of annotation results alongside other relevant parameters, such as demographic and ability measures, conditions, procedural steps, etc., providing the basis for statistical tests.
Triangulation and extensions	How do the identified patterns relate to other measures such as performance, reaction times, etc.? How can these results be used for further goals such as cognitive modeling, user support software, etc.?	CODA can readily be applied alongside other methods and frameworks. Triangulation means mutual validation of results. Results gained by CODA often feed easily into theoretical models or practical applications.

## 4. Discussion

As indicated above, CODA-related research traditions fall into two areas, namely the linguistic expression of mental representations and of complex cognitive processes such as problem-solving. Following the outline of the procedural aspects of the CODA methodology in Section 3, the contribution of CODA will now be discussed for each of these areas in turn.

### 4.1. MENTAL REPRESENTATIONS

As one prominent area relevant to the linguistic expression of mental representations, a wide range of publications emerged from the investigation of how speakers describe spatial scenes under various circumstances and settings (e.g., Carlson & Logan, 2001; Coventry, Carmichael, & Garrod, 1994; Gorniak & Roy, 2004). Taken together, these results show how the spatial setting interacts with the speakers' conceptualizations as well as with the current discourse task and discourse strategies, leading to systematic differences in linguistic representations depending on the variation of seemingly negligible factors. A general conclusion can be derived that whenever there is a difference in conceptualization (influenced by context factors), there will be a difference in linguistic expression if the speaker is free to verbalize the conceptual patterns (such as spatial relationships) without constraints. Such conceptual differences become apparent with any conceivable change in the discourse setting, be it the spatial configuration, the nature of the interaction partner (e.g., human or robot), or details of the task at hand. Furthermore, individual differences in the

ways in which a situation is perceived can lead to systematic differences in the verbalization data.

In this area of research, the linguistic structures that have been investigated using CODA have focused on principles of spatial term usage and their relationship to underlying spatial reference frames, perspective usage, and levels of detail or granularity. Relevant settings include descriptions of configurations in pictures (Tenbrink, 2007, Ch. 6; Tenbrink & Ragni, 2012; Vorweg & Tenbrink, 2007), complex configurations in small-scale space (Tenbrink, Coventry, & Andonova, 2011), and verbalizations of routes (Tenbrink, Bergmann, & Konieczny, 2011). The linguistic features that were analyzed to access conceptual distinctions in these settings primarily concerned the use of projective terms such as *left*, *right*, *in front of*, *behind*. Their use in relation to a specific configuration reflects the conceptual perspective adopted as well as the types of conceptualized object relation, as revealed by the choice of relatum (see Tenbrink, 2011, for the network of conceptual options available in this regard). Modifications and specifications of the spatial relation descriptions highlight the level of granularity perceived to be relevant by the speakers, and the order of mention reveals sequences and conceptual chunking of particular configurations (see also Plumert et al., 1995).

Clearly, the systematic analysis of unconstrained language poses a range of challenges. Allowing for participants' individual conceptualizations (and ensuing verbalizations) of the situation counters predictivity, and consequently may not always lead to statistically significant results. Moreover, it takes some amount of experience and subject knowledge to identify those structures in language that are relevant for a particular research issue at hand. Nevertheless, previous outcomes indicate that the effort is worthwhile, considering the broad range of specific results and publications emerging from this research. In general terms, the following contributions of linguistic analysis can be noted with respect to the study of mental representations.

- CODA brings together established research traditions in a productive way. Cognitive linguists have predominantly been concerned with generalized patterns of grammatical structures in language, in part focusing on cross-linguistic and cultural differences and their impact on thought. This research tradition only rarely incorporates the elicitation of language in controlled, cognitively interesting, settings in order to detect the situational parameters that lead to specific conceptual representations. Similarly, research in discourse analysis typically does not address cognitive issues in controlled situations for groups of speakers. This is achieved in psycholinguistic experimentation, which, however, rarely allows for unconstrained language use. CODA combines these three

areas by applying findings from cognitive linguistics to specific discourse settings that resemble psycholinguistic research scenarios, opening up further avenues for research in cognitive science based on established methodologies.

- The examination of linguistic features across a wide range of situations, taking into account the specific contextual features in each case, leads to a better understanding of the core semantics of specific lexemes. Thus, CODA-based analysis helps to distinguish those aspects of lexemes that remain unchanged across contexts (i.e., the semantic features) from those that are contextually variable (i.e., the pragmatic inferences that may be swiftly filled in by the discourse participants without conscious awareness). Insights along these lines for the spatial domain are represented in a spatial linguistic ontology by Bateman, Hois, Ross, and Tenbrink (2010).
- The findings obtained by linguistic analysis complement approaches from other research directions that also aim at a better understanding of mental representations. For instance, much research has dealt with the effects of relevance and salience on visual attention (Li, 2002; Navalpakkam & Itti, 2005; Ward, Duncan, & Shapiro, 1996). Where humans look at and focus on is guided, on the one hand, by outstanding distinguishing features of the perceived entities, and on the other hand by their current task purpose. These phenomena parallel findings reported with respect to language use quite closely; as predicted by Talmy (2007), the distribution of conceptual attention is systematically reflected in linguistic structures.

Linguistic analysis as described here thus amounts to a method for understanding mental representations that directly complements research in various fields. These parallel aspects can be addressed directly by relating other kinds of data to the results obtained by examining language. For instance, the investigation of eye-movements along with language production (Holsanova, 2008; van Gog, Paas, & van Merriënboer, 2005), the elicitation of memory data (Brunyé & Taylor, 2008), and the examination of gestures associated with language (Allen, 2003; Emmorey & Casey, 2002; Goldin-Meadow, 1999; Kranstedt, Lücking, Pfeiffer, Rieser, & Wachsmuth, 2006) provide relevant insights that complement the findings obtained by the analysis of unconstrained natural language data reflecting mental representations.

#### 4.2. COMPLEX COGNITIVE PROCESSES

The second relevant research tradition concerns the investigation of complex cognitive processes such as problem-solving via verbal protocols. To access human thoughts during problem-solving, behavioral tasks are often combined with some kind of linguistic description of the task and

its solution, allowing for a comparison of freely produced language with behavioral data. The analysis then highlights the use, function, and interplay of cognitive components and processes involved in the task as reflected jointly by language and behavior. CODA adopts this general approach as established by Ericsson and Simon (1993) and pursues it further by a closer examination of patterns in linguistic structure, in order to support the evidentiary value of verbal protocols as well as to highlight how particular conceptual aspects are reflected in language. Relating non-linguistic problem-solving behavior to linguistic strategies provides a strong motivation for adopting established goals and methods of linguistic discourse analysis. Furthermore, participants' utterances often contain explicit information about underlying strategies (e.g., Hölscher, Meilinger, Vrachliotis, Brösamle, & Knauff, 2006; Spiers & Maguire, 2008); such information is specifically interesting since speakers provide direct access to aspects that are important to them (Sacks, Schegloff, & Jefferson, 1974). Generally, as Caron (1996) puts it:

verbal protocols, not taken as descriptions of the subjects' mental processes, but as interpretable traces of those processes, can be a valuable source of on-line information about cognitive functioning. (Caron, 1996, p. 12)

One of the aspects that can be highlighted by language data concerns the identification and categorization of errors, mistakes, and false leads (e.g., Gralla, 2014). In complex problem-solving tasks, such failures may occur on different levels, leading to less than optimal results. According to Reason (1990), errors can be ascribed to specific subprocesses of human planning actions: for instance, either the intended plan itself (which is based on the participants' knowledge and inferential processes) is faulty, or the execution of the plan is other than intended. Purely behavioral results will not in all cases provide sufficient information to differentiate between these possibilities. Verbal data, on the other hand, provide further insights concerning the participants' underlying intentions and thus contribute to the understanding of the level at which errors occurred.

CODA-based research in the area of problem-solving so far includes planning paths to one goal location in an urban environment (Hölscher et al., 2011) and in a complex building (Tenbrink, Bergmann, & Konieczny, 2011), and to multiple goals in abstract configurations (Gralla et al., 2012; Tenbrink & Wiener, 2009) and in everyday contexts (Tenbrink & Seifert, 2011). Results encompass a broad variety of insights about speakers' meta-cognitive awareness of complex cognitive processes, a diversified repertory of conceptual strategies when addressing them, insights on conceptual focus and relevant granularity levels, as well as perspectives, dynamic shifts between conceptual domains, and a range of communicative aspects that mediate the verbalization of associated tasks and procedures.

The analysis of verbal protocol data using linguistic (discourse analytic) methods is as such a novel contribution to cognitive science. Discourse analysts have so far not been known to analyze language data related to cognitive science problems, and the systematic analysis of linguistic features is not an integral part of the widespread tradition using verbal reports as data, promoted most prominently by Ericsson and Simon (1993). In fact, Ericsson and Simon explicitly recommend transforming the original data to a generalized, abstract, form that is better suited for operationalization of annotation categories. While such a procedure is certainly both useful and common in content analysis, it leaves little room for the identification of informative linguistic patterns in the original data. For example, as outlined above, the close linguistic analysis by Tenbrink and Seifert (2011) highlighted how speakers swiftly and implicitly switched between the fundamentally distinct conceptual domains of planning and traveling; insights such as these could hardly be derived by way of generalizing statements from protocols. Furthermore, the specification of lexical and syntactic structures based on a close linguistic analysis can support insights about content. By focusing on a particular type of linguistic detail, it becomes possible to operationalize the analysis of specific aspects involved in a cognitive process. Instead of using broad content-based categories derived from the produced language, the analysis targets specific sets of linguistic features that have been identified as relevant indicators for particular issues.

The range of analysis avenues that can be chosen to investigate complex cognitive processes is naturally broader than for mental representations as described above. This is due to the higher amount of flexibility, i.e., the larger network of linguistic options, available to speakers in complex cognitive tasks (as opposed to the verbalization of a perceived scene). This raises the challenges to the analysis considerably, but also opens up a wider range of promising insights. Problem-solving tasks involve many different aspects that may become the target of linguistic analysis, including a timeline (which may be reflected by temporal discourse markers), conscious strategies and decisions (with lexical choices depending on the nature of the task), considerations of states of affairs (which, at each point in time, may be analyzed according to the mental representations involved), plans of possible future actions and caused states (which may be represented by modal verbs and causal connectors), conceptual changes caused by real-world actions and changes (which may be reflected by explicit linguistic signals of insight), mental switches between task domains (represented by different semantic fields), variations in attention focus (reflected in linguistic information structure and lexical choices), and many more. The cited CODA-based publications provide concrete examples along these lines; however, the range of insights that can potentially be gained by adopting a particular linguistic

analysis perspective is conceived to be far wider, leaving much room for future exploration.

Arguably, analyzing verbal data from a linguistic point of view is not just useful for addressing cognitive science issues, but actually a natural and necessary development emerging from accumulated progress on both sides – linguistic discourse analysis, and cognitive science. The following points support this view.

On the linguistics side, much debate has been devoted to the relationship between language and mind. A wide range of publications, many theoretical in nature, others supported by observing language(s) used in everyday life in different cultures, address the question of whether language determines thought – an idea that was formulated most pointedly by Whorf (1941). Other research directions address the conceptual structure represented by certain lexical items, particularly prepositions. Psychologists and psycholinguists have been extensively investigating the language used to describe mental representations (typically using a somewhat different approach than CODA). The step forward to investigate not only concepts and representations but also cognitive PROCESSES – such as those involved in problem-solving – is a direct and straightforward one. It sheds new light on the relationship between language and thought by offering new kinds of answers to old questions.

To be more precise, the long-standing question about the relationship between language and thought can be refined to encompass the following: To what extent does our language USE express what we think? Which kinds of linguistic structure reflect which kinds of cognitive process, how do they map, and how are they chosen from the available network of options? How do actions when solving problems affect language use? How can these insights be utilized for practical purposes, for example supporting complex cognitive demands by using appropriate language in instruction manuals as well as user support software? It is a fascinating prospect to use discourse analytic results to gain insights beyond the realm of linguistics, informing other strands of research related to cognitive science, and opening up a range of practical application aims such as intuitive human–robot and human–computer interaction (Mast & Bergmann, 2013; Moratz & Tenbrink, 2006).

Generally, this kind of language data offers an exciting new resource for texts that call for analysis by linguistics experts. Discourse analysis typically aims at a better understanding of how discourse works (which, undoubtedly, is an important aim in itself), particularly with respect to communication and (in Critical Discourse Analysis) with respect to manipulation. Verbal protocol data represent a fundamentally different text type (in comparison to everyday usage) that indicates how language may be used for a purpose that is not

primarily communicative. As observed by Caron-Pargue and Gillis (1996), the role of discourse markers, for instance, may change according to text type, ranging from a communicative function to that of signaling topic shifts and cognitive chunking processes. The analysis of linguistic reflections of conceptualizations of relations and entities, such as lexical choices on varying levels of granularity, in the course of problem-solving tasks, provides a better understanding of the referential scope and perceived prominence of the concepts involved.

In cognitive science, language data have been scrutinized to gain insights about cognitive processes for several decades. As part of these endeavors, specific analysis steps have been developed that can be reminiscent of established linguistic theories without being informed by them (based on the different scientific background of the researchers involved). To cite one example, Goldschmidt (1992, 2014) proposed a method for detecting and visualizing links between portions of linguistic protocols elicited by architects during design processes. The analogy to discourse analytic research on coherence relations within texts (Sanders, 1997), also called 'rhetorical structures' (Mann & Thompson, 1988), would be obvious to linguists, yet the approaches so far remain unrelated. Tenbrink (2008) provided a qualitative structural analysis as reflected by discourse markers (supported by relative frequencies); combining this approach with Goldschmidt's linkograph should provide a good basis for operationalizing and deriving further inferences from the analysis of coherence. Similar observations hold for other types of analysis carried out by cognitive scientists interested in systematic operationalizations of intuitively meaningful patterns found in linguistic data. Generally, the main benefit of systematic linguistic analysis arguably lies in the well-informed operationalization of coding categories based on established linguistic insights, supporting the validity of analysis results.

Moreover, linguistic structure can reveal patterns of thought (such as underlying spatial reference systems, focus of attention, granularity levels, conceptual perspectives, and the like) that speakers may not be consciously aware of and would not explicitly verbalize, thereby extending the limits of accessibility of cognitive representations to a degree. The limits of CODA-based research in this respect are clearly a matter of further exploration. While some reflection in language is naturally a prerequisite for any language-based analysis, specific analysis procedures may conceivably be further refined to allow for the identification of a wider range of unconscious cognitive processes than has been assumed so far.

Another way in which linguistic analysis can inspire research is to inform cognitive modeling efforts (e.g., using ACT-R; Anderson et al., 2004) by drawing on the prominent problem-solving steps as verbalized

by the participants, further specified with respect to underlying concepts by a close linguistic analysis. Various alternative solution paths may be available for the same task if participants differ in their cognitive strategies (e.g., Smith, Lewis, Howes, Chu, Green, & Vera, 2008). Typically, with increasingly complex tasks people have increasingly complex and flexible strategies at their disposal, which not only differ inter- but also intra-individually. Recognizing and specifying this flexibility to provide adequately versatile cognitive models is one of the prominent aims in cognitive science. Another aim is to capture the ways in which people develop their cognitive strategies in the first place, guided by learning procedures and (typically) analogical strategies, building on previous experience. These processes can be traced by think-aloud protocols spanning various instances of the same task.

## 5. Conclusion and outlook

This paper has outlined Cognitive Discourse Analysis as a method towards systematic analysis of unconstrained language data as evidence for cognitive processes and representations. This approach considers the cognitive linguistic repertory available to speakers to verbalize mental representations and complex cognitive processes. The analysis of linguistic patterns in speakers' unconstrained verbalizations highlights the relationship between language and thought in a particular context.

The methodology outlined here is based on a range of established and well-proven analysis procedures. It provides a coherent framework for the aim of accessing cognitive processes in easily accessible yet systematic ways. Many researchers may recognize considerable overlap with their own empirical designs, analysis methods, and procedural steps: this highlights the prominence and utility of such an approach, in spite of a predominant trend towards fine-grained experimental control, computational efficiency, and high-technology based procedures. Linguistic data collection, as such, is widespread and simple, requiring no specific technology (apart from standard audio recording devices) or expertise. In contrast to most other established methodological frameworks, however, the analysis of language for understanding cognitive processes so far has lacked a unifying concept, leading to frequent uncertainties and ad-hoc decisions whenever language comes into play. It is precisely the **UBIQUITY** of language within cognitive science and other areas of behavioral research that calls for a more rigorous approach. This paper has identified generic procedures, suggesting CODA – Cognitive Discourse Analysis – as a unifying term for approaches that use unconstrained language data to access cognitive processes and representations, across a broad range of research purposes.

## REFERENCES

- Afflerbach, P., & Johnston, P. (1984). On the use of verbal reports in reading research. *Journal of Reading Behavior*, **16**, 307–322.
- Allen, G. L. (2000). Principles and practices for communicating route knowledge. *Applied Cognitive Psychology*, **14**, 333–359.
- Allen, G. L. (2003). Gestures accompanying verbal route directions: Do they point to a new avenue for examining spatial representations? *Spatial Cognition and Computation*, **4**, 259–268.
- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, **111** (4), 1036–1060.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Barlow, M., & Kemmer, S. (Eds.) (2000). *Usage-based models of language*. Stanford, CA: CSLI Publications.
- Bateman, J., Hois, J., Ross, R. J., & Tenbrink, T. (2010). A linguistic ontology of space for natural language processing. *Artificial Intelligence*, **174**, 1027–1071.
- Bégoïn-Augereau, S., & Caron-Pargue, J. (2003). Linguistic criteria for demarcation and hierarchical organization of episodes in a problem solving task. In Frans H. van Eemeren, J. Anthony Blair, Charles A. Willard, & Francisca Snoeck Henkemans (Eds.), *5th Conference of the International Society for the Study of Argumentation* (pp. 81–87). Amsterdam: Sic Sat.
- Berman, R., & Slobin, D. I. (1994). *Relating events in a narrative*. Hillsdale, NJ: Lawrence Erlbaum.
- Best, B. J., & Simon, H. A. (2000). Simulating human performance on the traveling salesman problem. In N. Taatgen & J. Aasman (Eds.), *Proceedings of the Third International Conference on Cognitive Modeling* (pp. 42–49). Groningen: Universal Press.
- Biber, D. (1989). A typology of English texts. *Linguistics*, **27**, 3–43.
- Boroditsky, L. (2009). How does our language shape the way we think? In M. Brockman (Ed.), *What's next? Dispatches on the future of science*. Vintage Press.
- Boutonnet, B., Athanasopoulos, P., & Thierry, G. (2012). Unconscious effects of grammatical gender during object categorisation. *Brain Research*, **1479**, 72–79.
- Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, **34**, 383–398.
- Brösamle, M. (2013). *Sketches of wayfinding design: empirical studies of architectural design processes*. Unpublished doctoral dissertation, Albert-Ludwigs-Universität Freiburg i. Br.
- Brunyé, T. T., & Taylor, H. A. (2008). Working memory in developing and applying mental models from spatial descriptions. *Journal of Memory and Language*, **58**, 701–729.
- Cabeza, R., & Klingstone, A. (2001). *Handbook of functional neuroimaging of cognition*. Cambridge, MA: MIT Press.
- Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., & Anderson, A. (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics*, **23** (1), 13–32.
- Carlson, L., & Logan, G. D. (2001). Using spatial terms to select an object. *Memory & Cognition*, **29**, 883–892.
- Caron, J. (1996). Linguistic markers and cognitive operations. In J. Caron-Pargue & S. Gillis (Eds.), *Verbal production and problem solving* (pp. 11–28) (Antwerp Papers in Linguistics 85). Antwerp: Universiteit Antwerpen.
- Caron-Pargue, J., & Gillis, S. (Eds.) (1996). *Verbal production and problem solving* (Antwerp Papers in Linguistics 85). Antwerp: Universiteit Antwerpen.
- Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, **50**, 62–81.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**, 37–46.
- Coventry, K. R., Carmichael, R., & Garrod, S. C. (1994). Spatial prepositions, object-specific function and task requirements. *Journal of Semantics*, **11**, 289–309.

- Crampton, J. (1992). A cognitive analysis of wayfinding expertise. *Cartographica*, **29** (3/4), 46–65.
- Daniel, M.-P., & Denis, M. (2004). The production of route directions: investigating conditions that favor conciseness in spatial discourse. *Applied Cognitive Psychology*, **18** (1), 57–75.
- de Beaugrande, R. (1980). *Text, discourse and process: toward a multidisciplinary science of texts*. Norwood, NJ: Ablex.
- Degand, L., & Simon, A. C. (2009). On identifying basic discourse units in speech: theoretical and empirical issues. *Discours 4: Linearization and segmentation in discourse* (Special Issue), online: <<http://discours.revues.org/index5852.html>>.
- Denis, M. (1997). The description of routes: a cognitive approach to the production of spatial discourse. *Cahiers de Psychologie Cognitive*, **16** (4), 409–458.
- Ellis, A. W. (1985–1987). *Progress in the psychology of language*, 3 vols. London: Lawrence Erlbaum.
- Emmorey, K., & Casey, S. (2002). Gesture, thought, and spatial language. In K. R. Coventry & P. Olivier (Eds.), *Spatial language: cognitive and computational aspects* (pp. 87–101). Dordrecht: Kluwer.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: verbal reports as data*, 2nd ed.; first ed. 1984. Cambridge, MA: Bradford books/MIT Press.
- Evans, V. (2014). *The language myth: uncovering the true nature of language, mind and being human*. Cambridge: Cambridge University Press.
- Evans, V., & Green, M. (2006). *Cognitive linguistics: an introduction*. Edinburgh: Edinburgh University Press.
- Fincham, J. M., Carter, C. S., van Veen, V., Stenger, V. A., & Anderson, J. R. (2002). Neural mechanisms of planning: a computational analysis using event-related fMRI. *Proceedings of the National Academy of Sciences, USA*, **99**, 3346–3351.
- Findlay, J. M. (2004). Eye scanning and visual search. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: eye movements and the visual world* (pp. 135–159). New York: Psychology Press.
- Fischer, K. (Ed.). (2006). *Approaches to discourse particles*. Amsterdam: Elsevier.
- Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Science*, **3**, 419–429.
- Goldschmidt, G. (1992). Criteria for design evaluation: a process-oriented paradigm. In Y. E. Kalay (Ed.), *Evaluating and predicting design performance* (pp. 67–79). Chichester: J. Wiley.
- Goldschmidt, G. (2014). Linkography: unfolding the design process. Cambridge, MA: MIT Press.
- Gorniak, P., & Roy, D. (2004). Grounded semantic composition for visual scenes. *Journal of Artificial Intelligence Research*, **21**, 429–470.
- Graham, S. M., Joshi, A., & Pizlo, Z. (2000). The Traveling Salesman Problem: a hierarchical model. *Memory & Cognition*, **28**, 1191–1204.
- Gralla, L. (2014). *Linguistic representation of problem solving processes in unaided object assembly*. Unpublished doctoral dissertation, University of Bremen.
- Gralla, L., Tenbrink, T., Siebers, M., & Schmid, U. (2012). Analogical problem solving: insights from verbal reports. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 396–401). Austin, TX: Cognitive Science Society.
- Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions and the structure of discourse. *Computational Linguistics*, **12** (3), 175–204.
- Gugerty, L., & Rodes, W. (2007). A cognitive model of strategies for cardinal direction judgments. *Spatial Cognition and Computation*, **7** (2), 179–212.
- Halliday, M. A. K. (1994). *An introduction to functional grammar*, 2nd ed. London: Edward Arnold.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (1999) *Construing experience: a language-based approach to cognition*. London / New York: Continuum.
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, **1**, 77–89.

- Hayes-Roth, B., & Hayes-Roth, F. (1979). A cognitive model of planning. *Cognitive Science*, **3**, 275–310.
- Henderson, J. M., & Ferreira, F. (Eds.) (2004). *The interface of language, vision, and action: eye movements and the visual world*. New York: Psychology Press.
- Holsanova, J. (2008). *Discourse, vision, and cognition*. Amsterdam: Benjamins.
- Hölscher, C., Meilinger, T., Vrachliotis, G., Brösamle, M., & Knauff, M. (2006). Up the down staircase: wayfinding strategies and multi-level buildings. *Journal of Environmental Psychology*, **26** (4), 284–299.
- Hölscher, C., Tenbrink, T., & Wiener, J. (2011). Would you follow your own route description? *Cognition*, **121**, 228–247.
- Krahmer, E., & Ummelen, N. (2004). Thinking about thinking aloud: a comparison of two verbal protocols for usability testing. *IEEE Transactions on Professional Communication*, **47** (2), 105–117.
- Kranstedt, A., Lücking, A., Pfeiffer, T., Rieser, H., & Wachsmuth, I. (2006). Deictic object reference in task-oriented dialogue. In G. Rickheit & I. Wachsmuth (Eds.), *Situated communication* (pp. 155–207). Berlin: Mouton de Gruyter.
- Krippendorff, K. (2004). *Content analysis: an introduction to its methodology*, 2nd ed. London: Sage.
- Kuipers, B. J., Moskowitz, A. J., and Kassirer, J. P. (1988). Critical decisions under uncertainty: representation and structure. *Cognitive Science*, **12**, 177–210.
- Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 1–64). Stanford, CA: CSLI.
- Leidner, J., & Schilder, F. (2010). Hunting for the black swan: risk mining from text. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)* (pp. 54–59), Uppsala, Sweden, online: <<http://anthology.aclweb.org/P/P10/P10-4010.pdf>>.
- Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, **6** (1), 9–16.
- Lindsey, A. E., Greene, J. O., Parker, R. G., & Sassi, M. (1995). Effects of advance message formulation on message encoding: evidence of cognitively based hesitation in the production of multiple-goal messages. *Communication Quarterly*, **43** (3), 320–331.
- Mann, W. C., & Thompson, S. A. (1988). Rhetorical Structure Theory: toward a functional theory of text organization. *Text*, **8** (3), 243–281.
- Mast, V., & Bergmann, E. (2013). Is it really that simple? The complexity of object descriptions in human-computer interaction. In Markus Knauff, Natalie Sebanz, Michael Pauen, & Ipke Wachsmuth (Eds.), *35th annual meeting of the Cognitive Science Society: CogSci 2013* (pp. 990–995). Berlin, Germany.
- Miller, G. A. (1951). *Language and communication*. New York: McGraw-Hill.
- Moratz, R., & Tenbrink, T. (2006). Spatial reference in linguistic human-robot interaction: iterative, empirically supported development of a model of projective relations. *Spatial Cognition and Computation*, **6** (1), 63–106.
- Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research*, **45** (2), 205–231.
- Newcombe, N. S., & Ratliff, K. R. (2007). Explaining the development of spatial reorientation: modularity-plus-language versus the emergence of adaptive combination. In J. M. Plumert & J. P. Spencer (Eds.), *The emerging spatial mind* (pp. 53–76). New York: Oxford University Press.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Nuyts, J., & Pederson, E. (Eds.) (1997). *Language and conceptualization*. Cambridge: Cambridge University Press.
- Pick, H. L., Heinrichs, M. R., Montello, D. R., Smith, K., and Sullivan, C. N. (1995). Topographic map reading. In Peter A. Hancock, John M. Flach, Jeff Caird, & Kim J. Vicente (Eds.), *Local applications of the ecological approach to human-machine systems*, vol. 2 (pp. 255–284). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pinker, S. (1994). *The language instinct*. New York: William Morrow & Company.
- Plumert, J. M., Carswell, C., de Vet, K., & Ihrig, D. (1995). The content and organization of communication about object locations. *Journal of Memory and Language*, **34**, 477–498.

- Purcell, T., & Gero, J. S. (1998). Drawings and the design process: a review of protocol studies in design and other disciplines and related research in cognitive psychology. *Design Studies*, **19** (4), 389–430.
- Ragni, M., Fangmeier, T., & Brüssow, S. (2010). Deductive spatial reasoning: from neurological evidence to a cognitive model. In Dario D. Salvucci & Glenn Gunzelmann (Eds.), *Proceedings of the 10th International Conference on Cognitive Modeling* (pp. 193–198), online: <<http://iccm2010.cs.drexel.edu/proceedings/>>.
- Ranyard, R., Crozier, W. R., & Svenson, O. (Eds.) (1997). *Decision making: cognitive models and explanations*. London: Routledge.
- Reason, J. (1990). *Human error*. Cambridge: Cambridge University Press.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, **50**, 696–735.
- Sanders, T. (1997). Semantic and pragmatic sources of coherence: on the categorization of coherence relations in context. *Discourse Processes*, **24**, 119–147.
- Schelhorn, S. E., Griego, J., & Schmid, U. (2007). Transformational and derivational strategies in analogical problem solving. *Cognitive Processing*, **8**, 45–55.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.
- Schober, M. F. (1998). Different kinds of conversational perspective-taking. In S. R. Fussell & R. J. Kreuz (Eds.), *Social and cognitive psychological approaches to interpersonal communication* (pp. 145–174). Mahwah, NJ: Lawrence Erlbaum.
- Schober, M. F., & Brennan, S. E. (2003). Processes of interactive spoken discourse: the role of the partner. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 123–164). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: when language overshadows insight. *Journal of Experimental Psychology: General*, **122** (2), 166–183.
- Seifert, I. (2008). *Spatial planning assistance: a cooperative approach*. Unpublished doctoral dissertation, Universität Bremen.
- Selting, M. (2000). The construction of units in conversational talk. *Language in Society*, **29**, 477–517.
- Shi, H., Jian, C., & Rachuy, C. (2011). Evaluation of a unified dialogue model for human-computer interaction. *International Journal of Computational Linguistics and Applications*, **2** (1), 155–173.
- Smith, M. R., Lewis, R. L., Howes, A., Chu, A., Green, C., & Vera, A. (2008). More than 8,192 ways to skin a cat: modeling behavior in multidimensional strategy spaces. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1441–1446), Austin, TX.
- Sperber, D., & Wilson, D. (1986). *Relevance: communication and cognition*. Cambridge, MA: Harvard University Press.
- Spiers, H. J., & Maguire, E. A. (2008). The dynamic nature of cognition during wayfinding. *Journal of Environmental Psychology*, **28** (3), 232–249.
- Talmy, L. (2000). *Toward a cognitive semantics*, 2 vols. Cambridge, MA: MIT Press.
- Talmy, L. (2007). Attention phenomena. In D. Geeraerts & H. Cuyckens (Eds.), *Handbook of cognitive linguistics* (pp. 264–293). Oxford: Oxford University Press.
- Taylor, H. A., & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language*, **35**, 371–391.
- Tenbrink, T. (2007). *Space, time, and the use of language: an investigation of relationships*. Berlin: Mouton de Gruyter.
- Tenbrink, T. (2008). The verbalization of cognitive processes: thinking-aloud data and retrospective reports. In W. Ramm & C. Fabricius-Hansen (Eds.), *Linearisation and segmentation in discourse: multidisciplinary approaches to discourse (MAD 08)* (pp. 125–135). Lysebu, Oslo: Department of Literature, University of Oslo.
- Tenbrink, T. (2011). Reference frames of space and time in language. *Journal of Pragmatics*, **43** (3), 704–722.
- Tenbrink, T., Bergmann, E., & Konieczny, L. (2011). Wayfinding and description strategies in an unfamiliar complex building. In L. Carlson, C. Hölscher, & T. F. Shipley (Eds.),

- Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1262–1267). Austin, TX: Cognitive Science Society.
- Tenbrink, T., Coventry, K. R., & Andonova, E. (2011). Spatial strategies in the description of complex configurations. *Discourse Processes*, **48**, 237–266.
- Tenbrink, T., & Ragni, M. (2012). Linguistic principles for spatial relational reasoning. In C. Stachniss, K. Schill, & D. Uttal (Eds.), *Spatial Cognition 2012* (LNAI 7463) (pp. 279–298). Heidelberg: Springer.
- Tenbrink, T., Ross, R. J., Thomas, K. E., Dethlefs, N., & Andonova, E. (2010). Route instructions in map-based human-human and human-computer dialogue: a comparative analysis. *Journal of Visual Languages and Computing*, **21** (5), 292–309.
- Tenbrink, T., & Seifert, I. (2011). Conceptual layers and strategies in tour planning. *Cognitive Processing*, **12** (1), 109–125.
- Tenbrink, T., & Wiener, J. (2009). The verbalization of multiple strategies in a variant of the traveling salesperson problem. *Cognitive Processing*, **10** (2), 143–161.
- Thomas, L. E., & Lleras, A. (2007). Moving eyes and moving thought: on the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin & Review*, **14**, 663–668.
- Tomasello, M. (2003). *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tversky, B. (1999). Spatial perspective in descriptions. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 109–169). Cambridge, MA: MIT Press.
- van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & Society*, **4** (2), 249–283.
- van Gog, T., Paas, F., & van Merriënboer, J. J. G. (2005). Uncovering expertise-related differences in troubleshooting performance: combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology*, **19**, 205–221.
- Vorweg, C., & Tenbrink, T. (2007). Discourse factors influencing spatial descriptions in English and German. In T. Barkowsky, M. Knauff, G. Ligozat, & D. Montello (Eds.), *Spatial cognition V: reasoning, action, interaction* (pp. 470–488). Berlin: Springer.
- Ward, R., Duncan, J., & Shapiro, K. (1996). The slow time-course of visual attention. *Cognitive Psychology*, **30**, 79–109.
- Whorf, B. L. (1941). The relation of habitual thought and behavior to language. In L. Spier (Ed.), *Language, culture, and personality: essays in memory of Edward Sapir* (pp. 75–93). Sapir Memorial Publication Fund: Menasha, WI.

## APPENDIX

Training task for think-aloud procedure (2–3 min.) adapted from Ericsson and Simon (1993). The instruction to participants should be conveyed orally, as close to the following as possible:

*In our study we are interested in what you think about as you perform a task that we give you. In order to do this I am going to ask you to THINK ALOUD during the whole procedure of the task. That is, I want you to say EVERYTHING you are thinking from start to finish of the task. I would like you to talk aloud CONSTANTLY. Don't try to plan out what you say and don't talk to ME. Just act as if you were speaking to yourself. It is most important that you keep talking, even though you won't get any response or feedback. Do you understand what I want you to do?*

*Good, now we will begin with some practice problems. First, I want you to multiply two numbers in your head and say out loud what you are thinking as you get an answer.*

*'What is the result of multiplying  $24 \times 36$ ?'*

*Good. Any questions? – Here's your next practice problem:*

*'How many windows are there in a house you grew up in?'*

After explaining the task procedure, the following instruction can be added:

*Don't forget to THINK ALOUD while doing so.*

*I won't interrupt you, and I won't judge your decisions and thoughts. We are interested in your thoughts while you do the task.*

During the task the experimenter makes sure that the participant keeps thinking aloud, and reminds them to do so if they forget or fall silent for more than a minute or so. Reminders need to be kept neutral, as in *Keep talking; keep thinking aloud.*