

Title

The Foundation for a Dialogic Grammar

By

Ivar Tormod Berg Ørstavik

The Faculty of Informatics and e-Learning

Sør-Trøndelag University College

NO-7004 Trondheim

Norway

Email: ivar.orstavik@hist.no

Phone: +47 92030838

Abstract

This article discusses what empirical foundation is needed to construct a dialogic grammar. When building grammars, linguists first set up a “data base” that can contain the language data relevant for its understanding of linguistic/discourse structure (Givon 1979). This article analyzes the established language data bases that underlie mainstream linguistics and conversation analysis, and then discusses: what language data base a dialogic grammar needs; the problems of setting up such a data base, and the methodological and empirical restrictions of such a language data base.

The Foundation for a Dialogic Grammar

1.1 Intertextuality and the goal of a dialogic grammar

The emergence of the internet has pronounced the existence of intertextuality for all to see. Hyper-text links make intertextual connections explicit both within and across texts, words, images, videos, etc. Similarly, chat rooms and email exchanges enable electronic conversations to reference, merge into, and branch out of each other. Even the computer programs making all this happen are simply texts of symbols and words that explicitly link functions in different programs together (Ørstavik 2008). Thus, the internet and ICT present intertextual structures as simple and apparent, and intertextuality is no longer seen as an exotic literary phenomenon but as an important part of everyday communication.

As the use of explicit intertextual links has increased, so has their systematic analysis. One prominent example of intertextual analysis is the PageRank algorithm in the Google search engine (Wikipedia 2008). The PageRank algorithm systematically describes hyper-text links between internet texts and calculates a score for every text based on who links to whom. A second example is email spam filters. Spam filters may stop emails that contain banned words or are sent from email addresses from which the recipient does not wish to hear. By systematically identifying intertextual patterns within and across email exchanges, spam filters can build hypotheses about whether an email is relevant for its users or not. A third example is the execution of computer programs. To execute or run a computer program, the machine responsible attempts to interpret the words and symbols in the code as references to functionality and data. This functionality and data are in turn kept in other program texts referencing yet other texts. To execute a computer program is essentially to run through interpretations of intertextual links to and from a potentially vast, coherent web of different program texts. The PageRank algorithm, spam filters, and computer programs all show how systematic analysis of intertextual structures is already put to practical use.

In academia, the study of language structures has been the domain of grammar. However, as this study will show, studies in grammar have largely focused on language structures within sentences and conversations and have not attempted to systematically describe structures beyond these borders, such as intertextuality. As a response to this gap and the need for systematic analysis of intertextual structures both in practical applications (described above) and research (Linell 2005), this study will undertake a new type of grammatical investigation: to combine dialogism and its approach to intertextuality with grammar.

1.2 The paradoxes surrounding a dialogic grammar

Theoretically, combining dialogism with grammar is a paradoxical venture. Opposition and even outright antagonism towards mainstream linguistics is a founding tenet in dialogism (Bakhtin 1986: 68; Holquist 1990: 44pp). Grammar construction, on the other hand, is considered the true child of mainstream linguistics, and so to build grammars is perceived as doing mainstream linguistics. Thus, combining the two risks corrupting one or both: the original grammarian's stance can be thwarted by dialogic criticism to suit its purposes, establishing one or more dialogic grammars that are grammars in name only; or the original dialogic opposition can be ignored to nurture the mainstream linguistic project of building grammars, comparable to brooding a cuckoo's egg. If not resolved, pursuing the paradox of a dialogic grammar is neither in the interest of dialogism nor grammarians.

However, this paradox rests upon several assumptions that need verification. First, the assumption that grammar is synonymous with mainstream linguistic is true only in the sense that so far, most grammars have been built by mainstream linguists. But established tradition does not dictate future work. Hence, if a grammarian's approach can be established independently from mainstream linguistics, then grammar might be disentangled from the primarily theoretical conflict between dialogism and mainstream linguistics. Second, dialogism intrinsically opposes any and all holistic, monologic language systems. Compared to other theories of language, dialogism approaches a chaotic phenomenon by

being somewhat chaotic itself. However, a systematic approach does not necessitate a holistic, monologic view of the phenomenon studied: it is neither impossible nor difficult to describe a phenomenon systematically without claiming to capture it entirely. Building grammatical systems requires only an internal holism, not necessarily an external holism, towards the language it describes. Thus, the intrinsic resistance towards language as a holistic, monologic phenomenon should not be equated with a resistance towards the use of holistic, monologic systems as parts of a wider, dialogical picture.

What then if the paradoxes of a dialogic grammar can be resolved? The possibilities of such a scenario are intriguing. First, a dialogic grammar would respond to the many requests for more systematic, analytical methods that adhere to dialogic theory (Linell 2005). Such a tool would be a valuable contribution to both researchers and practitioners. Second, untangling grammar from mainstream linguistics is in itself a useful exercise that extends our understanding of grammar. Synthesizing this independent concept of grammar with dialogism will expand this insight even further. Third and last, even the pursuit of a dialogic grammar in itself is a valuable experiment. In dialogism, several ideas have been theoretically deliberated, but approaching the same ideas methodologically as grammatical concepts calls for new debates.

1.3 The problem of separating grammar from linguistics

Alternative grammars based on alternative theoretical approaches to language have been proposed in the past. One such group of grammars, called 'functional grammars' attempts to forge new grammars within socially-oriented studies of discourse and language. These grammars illustrate how difficult it can be to separate grammar from mainstream linguistics. Nichols (1984: 99, 103) points out that several functional grammars share some internal contradictions: "[Reasons] for calling an analysis 'functional' [...] sometimes use a different sense than that actually employed in the analysis itself", and that "not infrequently an analysis will be form-to-function [despite claims] that form is dependent on function". Such discrepancies suggest that the new grammars carry premises about language and

discourse that surface in the actual analysis and that turn out to clash with premises postulated in the surrounding theory. Somehow, implied premises from other linguistic disciplines have unintentionally been smuggled into this new theoretical domain along with the concept of grammar.

To avoid such internal contradictions in a new dialogic grammar, this study aims to isolate an understanding of grammar from linguistic ideas about language. To make this distinction, however, is hard since theoretical ideas about language often are implicitly presented as intrinsic to grammars. In order to free our understanding of grammar from familiar associations, the existing (or potential) foundation of grammars in three different linguistic disciplines are analyzed: mainstream linguistics, conversation analysis and dialogism.

This interdisciplinary scope has an important upside: juxtaposing grammars from several disciplines forms a background against which differences and similarities are likely to stand out. Seeing these differences and similarities is what is needed to establish a more independent understanding of grammar. But the interdisciplinary scope also has a downside: a wide focus yields a superficial and partial analysis of each discipline. Readers are likely to perceive the analysis as not giving their discipline's depth and diversity the attention it is due. I shall return to this issue in the conclusion.

The study is conducted as follows: first, a short discussion of the term and concept of grammar is presented. Second, the theoretical foundation behind mainstream linguistic grammars is presented based on Givon's (1979) and Itkonen's (1996) related critiques. This presentation results in a description of the "language data base" (Givon 1979). The language data base makes up the empirical platform upon which grammars are built and can be described as a discipline's compromise between different theoretical and methodological concerns. Third, the language data base within conversation analysis is analyzed. This analysis illustrates two core aspects of grammars: 1) how different theoretical and methodological approaches establish different language data bases and thus

different grammars; and 2) how different grammars behave similarly within the confines of their language data bases. Fourth, two theoretical models from dialogism are discussed as theoretical premises for a dialogic grammar. This comparison illustrates the key problem of how to construct a unitary language data base in dialogism and how this problem might be addressed using the diatope model (Evensen 2002). The study concludes with some thoughts about future work towards a dialogic grammar.

2.1 What is grammar?

To define or even adequately describe "what grammar is?" is problematic. The first problem is that grammars and grammar-like concepts are at work in highly different disciplines and languages. Grammars are concepts associated with particular languages such as English and Norwegian, different branches of linguistics such as Chomskian linguistics and conversation analysis, formal languages such as mathematics and computer programming, to name a few. A description of grammar suitable in one of these frameworks might clash with the framework of the others. The second problem is that as these disciplines evolve and diversify, so do their grammars. Any description of grammar is therefore likely to be highly context sensitive, that is, fitted to a particular research group, purpose, and time. As the project of a dialogic grammar attempts to draw upon insight from several different disciplines in order to explore a new grammar, this article therefore faces both the problems of the diversity of existing grammatical concepts as well as the problems of fitting a new grammar to a new academic context.

As a term, 'grammar' is commonly used to refer to some kind of abstract, systematic description of rules or conventions in a language. However, as soon as one starts to pursue questions such as "abstracted how?", "what kind of systematicity?", "where do the rules or conventions come from?", and "what is meant by 'language' and 'in language'?", many different answers are given. These questions address the foundation of grammars, and as these foundations vary, so will the resulting description of grammars.

For example, the origin of grammar are approached as: (a) cognitive in generative linguistics, (b) social in sociolinguistics, (c) historical in comparative philology, (d) utilitarian in computer programming, and (e) Newtonian laws in mathematics. The differences with regards to the origin of language systems should not be understood as true or false. These differences vary according to the purposes and settings of the different grammars and inform each other. Furthermore, the concepts of 'abstraction' also differ. In philology and sociolinguistics, for example, grammars are post-priori constructs: descriptions of patterns in observed oral or written communication. In generative linguistics and computer programming, on the other hand, grammars are a-priori constructs: a theoretical hypothesis that is to be tested empirically or a normative tool and rules that prescribe how computer programs should be written. In practice, all of these disciplines go back and forth between their grammatical constructs and empirical basis; however, in theory, they see their grammars as abstractions that are either extracted from or projected onto empirical data. Lastly, the forms of systematicity also differ between these traditions. Some grammar and grammar-like constructs follow rigid and holistic formalisms; other grammars are less rigidly defined and deliberately describe only parts of their language domain. To conclude, a set of questions support the concept of grammar, and by pursuing the different answers to these questions, an understanding of "what grammar is?" can be established relative to different research contexts.

To develop a new dialogic grammar, this article will begin the examination of one such fundamental question: what is the empirical basis for different grammars, that is, what do different linguistic disciplines mean by 'language' and 'in language'? Based on the discussion of this question, a second question is posed: what kind of empirical basis do we need for a dialogic grammar, and how can grammars in other linguistic disciplines inform the concept of a dialogic grammar? To address these questions is an important first step and premise for working towards a new dialogic grammar. However, an answer to these questions by no means completes a dialogic grammar. Other important questions regarding systematicity, abstraction, and origin remain open and in need for future exploration.

2.2 The language data base

A grammar is a system built to describe ‘a language’, and thus, to define ‘what that language is’ precedes any grammars. Before linguists can construct or test their grammars, they must first stipulate what language data the grammar will describe. Different linguistic disciplines have different answers to ‘what language is’, and so different disciplines end up framing what Givon (1979) terms different **language data bases**, which again, result in strikingly different grammars.

As will become evident, theoretical principles are not the only factors that influence the make-up of language data bases. Methodological and practical concerns regarding how and what language data are collected also play an important role. But regardless of why each discipline chooses to include and exclude different types of language data, the form of the resulting language data base is crucial for the quality of the resulting grammar: “one of the most damaging practices in the history of linguistics has been the arbitrary – and a priori – delimitation of the data base, that is, the range of facts that serve as input for the investigation and, ultimately, as input for the building of the theory” (Givon 1979: 3). The demarcations of language data bases¹ must therefore be analyzed without prejudice.

2.3 Mainstream linguistics: the sentence data base

Chomsky (1969: 57) delimitates the language data base of modern linguistics as “an infinite set of sentences”: a sentence data base (as cited in Itkonen 1996: 472). Upon this empirical platform, a grammar that systematically describes humans’ intrinsic and mental structures that generate and interpret language is pursued (Chomsky 1965: 4). The result is “a generative grammar [simply as] a system of rules that in some explicit and well-defined way assigns structural description to sentences” (ibid.: 8).

¹ The term ‘data base’ refers here to an abstract entity, not computer systems for storing data. A language data base may well be set up using a computer system for storing data, but a language data base can also be set up using a researcher’s personal memory. A ‘language data base’ refers to the collection of ‘empirical language data’ from/against which a grammar is abstracted/contrasted.

The **sentence data base** has one very important *enabling* restriction. Grammatical structures can be determined based on ‘ideal sentences’: sentences intuitively abstracted from their context of use and isolated from each other one-by-one. Examining ideal sentences has many advantages. First, it provides a clear theoretical focus on internal sentence structures. This clarity is especially useful when these structures are part of a complex context of use or intertwined with each other: if, for example, internal sentence structures are to be analyzed in utterances stemming from a heated debate, the context of use might confuse the analysis. As long as it is *only* the internal structures that are sought, excluding relationships and context outside the sentence is beneficial. Second, ideal sentences are easy to come by. Sentences isolated from both each other and their context of use can be produced intuitively by the linguist himself. Also, the task of transcribing and cataloging the data base presents few challenges. Such easy access to data provides clear methodological and practical advantages.

However, enabling grammatical structures to be assigned to sentences one-by-one also has the negative implication that all relationships between – and thus requiring – two or more sentences together fall outside the frame of the sentence data base and thus outside the scope of the grammar. This precludes both the sentence data base and grammars built upon it from describing any linguistic phenomena that include relationships beyond the ‘ideal sentence’. To understand the consequences of these restrictions, we will now look at conversation analysis and its language data base and approach to grammar.

2.4 Conversation analysis: the conversation data base

While mainstream linguistics systematizes internal sentence structures, conversation analysis (CA) systematizes internal structures in talk-in-interaction. The focus of CA is the sequence and relationships between *turns* in conversations. As a language data entity, the turn can resemble the sentence in size, but apart from size, the two have little in common. While the sentence is abstracted from its context of use, the turn is always viewed as part of

a conversation: a turn is an utterance spoken by a particular (social) actor, in a particular situation, and at a particular time. Relevant first and foremost in CA are these links between the utterance and these contextual factors, not so much relationships within the utterance. In addition, turns are not isolated one-by-one but interconnected one-by-one like teeth in a zipper. So while mainstream linguistics describes a word sequence within individual sentences, CA describes turn sequences or turn-taking within individual conversations. The result is a **conversation data base** – ‘a set of all possible conversations’.

Even more so than mainstream linguistics, CA illustrates the methodological influence on the definition of language data base. CA was created “on the basis of research using audio recordings of naturally occurring conversations” (Sacks 1974: 697). Audio, and later video, recording made it practically feasible to collect, replay, and thus transcribe individual conversations as language data with sufficient detail and accuracy. Audio tape technology, as well as the method of conversation transcription that evolved with it, was a necessary precondition for the conversation data base.

On top of the conversation data base, CA “attempt[s] to characterize, in its simplest systematic form, the organization of turn-taking for conversation” (ibid.). This turn-taking system consists of a set of rules for how turns are recognized and ordered. These rules describe *inter*-turn structures, such as turn allocation, turn exchanges, adjacency pairs, reparation, etc. – structures that are *internal* in each conversation (ibid.; Hutchby 2008). As such, the turn-taking system can be understood as a conversation grammar: “a system of [conventions] that in some explicit and well-defined way assign structural descriptions to” conversations (cf. Chomsky 1965: 8).

The conversation data base spans a wider social and textual space, compared to the sentence data base, thereby enabling grammatical descriptions of interactional structures. But the conversation data base is still too restrictive to be a platform for systematic descriptions of several well-known socio-cultural linguistic relationships. The most pertinent restriction for dialogism is the conversation data base’s focus on isolated

conversations. Turns are depicted as dependent upon each other within each conversation but never outside the boundaries of individual conversation. Due to the amount of labor needed to record and transcribe one conversation, CA does not require more than one conversation for its analysis. As long as the systematic descriptions do not aim to describe structures beyond isolated conversations, these methodological and practical concerns can be headed. But if and when attempts are made to systematically describe relationships and structures, such as intertextuality, that span beyond isolated conversations, then the implied premises of the conversation data base become problematic.

To sum up, comparing the sentence and conversation data bases has illustrated four subtle fundamental premises for grammars. First, different theoretical views of language prescribe different language data bases. Second, methods and practical concerns also play an important role in shaping a language data base. Third, different language data bases form different grammars. And fourth, despite their differences, grammars resemble each other. Grammars describe the sequential structures between sub-entities, such as words or turns, structures that are internal to their primary language data entity, such as sentences or conversations. In this next section, I will use these four criteria to develop a language data base to underlie a dialogic grammar.

3.1 Demarcations of language data in dialogism

In one way or another, all theoretical discussions of language distinguish between relevant and irrelevant language data. Hence, several demarcations of language data bases have been put forth in dialogism, some overlapping and some complementing each other. So even though examining the foundation for a dialogic grammar might be a new approach within dialogism, examining the contours of a language data base is not.

Bakhtin (1981: 276) gives this description of words and their environment.

[N]o living word relates to its object in a singular way: between the word and its object, between the word and the speaking subject, there exists an elastic environment of other alien words about the same object, the same theme, and this is an environment that is often difficult to penetrate. It is precisely in the process of living interaction with the specific environment that the word may be individualized and given stylistic shape.

Indeed, any concrete discourse (utterance) finds the object at which it was directed already as it were overlain with qualifications, open to dispute, charged with value, already enveloped in an obscuring mist – or, on the contrary, by the ‘light’ of alien words that have already been spoken about it. The word is entangled, shot through with shared thoughts, points of view, alien value judgments and accents. The word, directed towards its object, enters a dialogically agitated and tension-filled environment of alien words, value judgments and accents, weaves in and out of complex interrelationships, merges with some, recoils from others, intersects with yet a third group: and all this may crucially shape discourse, may leave a trace in all its semantic layers, may complicate its expression and influence the entire stylistic profile.

The living utterance, having taken meaning and shape at a particular historical moment in a socially specific environment, cannot fail to brush up against thousands of living dialogical threads, woven by social-ideological consciousness around the given object of an utterance; it cannot fail to become an active participant in social dialogue. After all, the utterance arises out of this dialogue as a continuation of it and as a rejoinder to it – it does not approach the object from the sidelines.

The way in which the word conceptualizes its object is a complex act – all objects, open to dispute and overlain as they are with qualifications, are from one side highlighted while from the other side dimmed by heteroglot social opinion, by an alien word about them. And into this complex play of light and shadow the word enters – it becomes saturated with this play, and must determine within it the boundaries of its own semantic and stylistic contours.

Put bluntly, words are related to each other both within and across “a thousand” dialogues. These relationships merge, recoil, and intersect words with each other, and these intra- and inter-dialogical relationships between words are essential for their semantic and structural form. “Word” refers here to instances of uttered words, but the status of the word is not elevated above other linguistic entities, and similar interpretations for relationships between phrases, utterances, whole texts, etc. are not excluded.

However, the texts of the Bakhtinian circle are too diverse to simply be ‘translated’ into a language data base definition. Even translations of individual terms, such as ‘word’ and ‘heteroglossia’, from Russian to English are disputed, and so to simply forge a language data base from these original texts is unlikely to yield any demarcation viable either theoretically or methodologically. Therefore, contemporary studies from the Scandinavian school of applied dialogism have presented theoretical models that are considered largely in line with dialogic theory. This school analyzes written and verbal communication and thus needs a reliable and verifiable empirical platform to perform its analysis. Questions such as “what language data is needed?” and “how do these data correspond to dialogic theory?” have already been asked, and this study builds on and within this school’s previous work.

3.2 Problems with the double dialogue as a language data base

Linell (1998) uses the term “double dialogue” to describe the scope of language data in dialogic theory. This term emphasizes two simultaneous dimensions: here-and-now interaction and historical constructionism. First, dialogue should be understood literally in reference to actual conversations, novels, email exchanges, and other situations in which people interact with language. This here-and-now dialogue echoes the conversation in CA. But dialogue means more than just here-and-now conversations: dialogue also refers to a second, socio-historical dimension. Words, utterances, and even here-and-now dialogues

themselves are “in dialogue” with each other over a longer historical timeframe². These cross-dialogue dialogues shape how people and cultures socially construct such things as a context and semantic background, which is instrumental in making utterances in here-and-now dialogues meaningful.

Theoretically, the double dialogue model illustrates a view of a language data base compatible with dialogism. However, despite the simplicity of the double dialogue model, the model also suggests a duality: the dialogic relationships within a conversation are qualitatively different from the dialogic relationships between two or more conversations. This duality can be interpreted to mean that dialogic relationships should be viewed in two intersecting, but still separate, language data bases: one for dialogical relationships within individual here-and-now dialogues, and another for dialogical relationships between here-and-now dialogues.

But using two language data bases as the foundation for one grammar is highly problematic. As the previous analysis of mainstream linguistic and CA illustrated, existing grammars seek internal structures within *one* language data entity, embodied as *one* language data base. If, on the one hand, the double dialogue model theoretically stipulates that language data are dialogical relationships both within and between here-and-now dialogues, then a dialogic grammar building upon this model would have as its *one* language data entity individual clusters of here-and-now dialogues, not individual here-and-now dialogues. But if, on the other hand, language data are captured in two separate dimensions, structures sought within the one language data entity are to be found methodologically spread across *two* data bases. Hence, a system that can theoretically be described as within the double dialogue is thus methodologically internal to neither of the two language data bases, but rather resides both within and between the two. To develop any system based on such a complex data basis that is coherent enough to be considered a grammar seems both an insurmountable task and in conflict with the established practice of “one grammar, one language data base”.

² cf. “thousands of living dialogical threads” (Bakhtin 1981: 276)

Previous works on a “dialogical grammar” (Linell 2005; Norén 2007) should be understood against this platform. Methodologically, these studies are based in CA: “Interactional linguists and dialogical grammarians make use of fine-grained analysis of speakers’ activities at both the single utterance level and the sequential level in the spirit of conversation analysis (CA)” (ibid.: 10). But theoretically, “dialogical grammar” is at least partly positioned within applied dialogism (cf. Linell 1998). However, when dialogic theory is combined with CA methods as the background for a “dialogical grammar”, then the conversation data base is imported without debate. No second language data base is suggested for relationships across dialogues. This compromise avoids the pitfalls of building a dialogic grammar based on two language data bases, but at the same time, it essentially casts the dialogical grammar as a “single dialogue grammar”, that only in theoretical underpinning distinguishes itself from existing conversation grammars based on the conversation data base.

3.3 Uniting the double dialogue in the diatope data base

“[T]he notion of dialogism has a much broader scope than that of actual dialogue. It goes beyond the direct cooperation between interlocutors in a given speech situation.” (Gasparov 2008: 26). According to dialogic theory, utterances are in dialogue both internally and externally – with other utterances in the here-and-now dialogue and with other utterances in past, parallel, and future dialogues *at one and the same time*. Furthermore, these relationships influence each other: for example, the use of a word in one dialogue may reciprocally affect the use of the same or another word in another dialogue. This complex interplay between both inter-dialogic and intra-dialogic structures (cf. intertextual and intra-textual) is the *double* in the double dialogue and the focus of recent dialogic theory.

In an effort to unite the intertextual and intra-textual structures within a single theoretical model, Evensen describes the domain of language data as a “diatope” (Evensen 2002). The diatope, simply meaning the space of dialogues, can be described as a single image that

encompasses intra-textual relationships, linking utterances within a single dialogue, as well as intertextual relationships, linking utterances in two or more different dialogues.

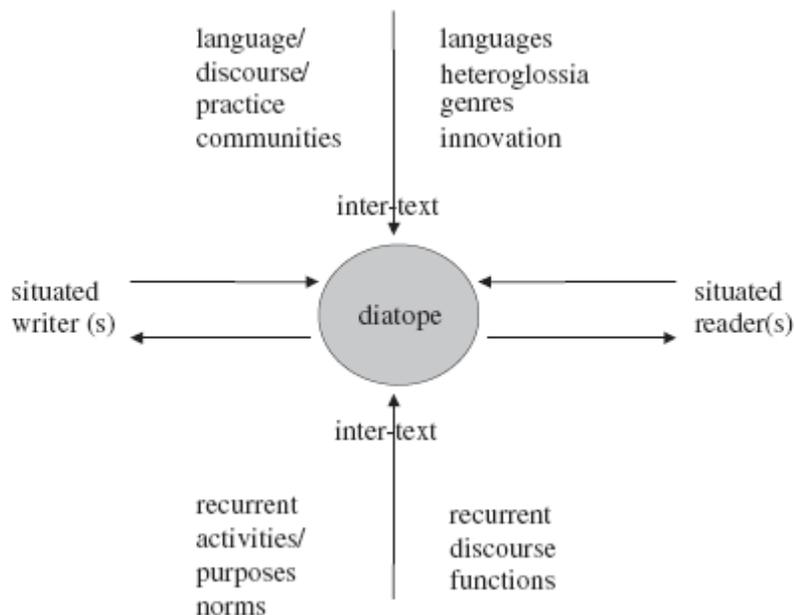


Figure 1: The diatope model (*ibid.*)

By interpreting the diatope as a description of a language data base, then the language data entity that contains the structures being pursued would be a body of utterances found across several dialogues/texts. This body of utterances resembles the body of sentences in corpus linguistics. But unlike a traditional text corpus, the social and situational position of each utterance would need to be included; for example, CA transcribes the speaking subject, time, and place. These social and situational anchors are necessary to order utterances both within each dialogue/text and between different dialogues/texts; without social and situational anchors for each word/text, it would be impossible to keep them apart both within and across each dialogue/text and, hence, impossible to link them together. The diatope data base must be ‘an unlimited set of clusters of dialogues with interrelated utterances’.

The main advantage of the diatope data base is its simple, unitary structure of utterances: once collected, the language data need not be stored in any special structures or hierarchies. However, the task of collecting language data to validly populate the diatope data base presents different challenges for different domains.

In computer-mediated written discourse, it is both technically and practically feasible to gather and transcribe enough utterances to build diatope data bases. Web pages, chat rooms, email, computer programs, etc. are readily available for existing methods of data collection and can relay enough intra+intertextual structure to form a diatope data base valid for partial intertextual analysis. The PageRank algorithm is a good example of the potential value of even very narrow systematic descriptions of intertextual structures (Wikipedia 2008). Principally, the intertextual, dialogical data base underlying the PageRank algorithm contains information only about which web pages link to what other web pages. These links function as a transcription of a socio-textual hierarchy modeled after similar hierarchies in academic papers. By simply rating prominent web pages by how many links they receive from other prominent web pages recursively, the PageRank algorithm forms a very small socio-textual grammar based on a quantitatively large, but qualitatively small, diatope data base.

In face-to-face discourse, however, collecting enough language data to construct a valid diatope might be more difficult using existing methods. For example, setting up a diatope data base focusing on 'coffee-break conversations at a workplace', one would face several obstacles. At the outset, the task of gathering and transcribing language data from real conversations is not yet fully automated, is therefore very costly, and often involves practical and ethical problems. This greatly restrains the researchers who must limit their diatope data base to, for example, ten socially related conversations. However, despite a small diatope data base, it might still be possible to find dialogic patterns across different conversations. Theoretically, dialogism gives no instruction to only describe local, sentence-like patterns in isolated conversations in grammatical terms and exclude patterns spanning longer chains of utterances between several socially related conversations. On the

contrary, if systematic patterns can be found and described between utterances across several socially related conversations, such descriptions will not only extend our understanding of (face-to-face) communication, but also potentially ground dialogic theories about polyphony and intertextuality empirically.

4.1 Future directions

Dialogism is a fundamental philosophy, and to constrain dialogic theory to so-called ‘natural language’ discourse is as restrictive to the enhancement of dialogic theory and dialogic methods as attempts to constrain linguistic research to the study of sentences only. Human discourse is rich and diverse, not only within the paradigm of natural language, but also outside this paradigm. Continuously expanding our studies of dialogues into such strange domains of human discourse will not only broaden our understanding of ‘what language *can* be’ (as opposed to ‘what language is’), but also broaden our understanding of what dialogic theory and methods *can* be.

This ‘un-natural’ pursuit of language has been the empirical cornerstone of this study. The diatope data base presented here was developed in an ongoing project to build a dialogic grammar for computer programming language discourse. Parallel to any other form of human written discourse, computer programmers write texts that have intertextual relationships with other computer programmers’ texts, both internally in here-and-now social situations and across such situations. A dialogic grammar would be strategic both in terms of understanding these dialogical relationships theoretically and in developing practical solutions that empower individual programmers in discourse. Computer programming is also an area of human discourse relatively isolated from outside influences and already transcribed, and therefore is an example of a domain in which the methodological and practical restrictions of the diatope data base can be addressed.

Future studies of a dialogic grammar can also draw some lessons from this study’s interdisciplinary method of extracting and contrasting grammar concepts from other related

linguistic discipline. Viewed negatively, this study's interdisciplinary analysis can be perceived as a caricature of the disciplines and their projects. First of all, the diversity of neither linguistic nor conversation studies is presented. Second, both linguistic and conversation studies are painted with the words of the other discipline: the study of grammars is presented as growing out of an empirical foundation, which goes in the opposite direction to its own internal order between theory and empirical data; and the study of conversations is linked to a grammatical goal, which is almost a contradiction in terms of its own project. Thus, if one seeks a comprehensive review that compares each tradition on its own premises, this interdisciplinary method is not the answer.

However, viewed positively, the interdisciplinary analysis extracts one important trend from each discipline. This article claims that these trends are in fact good, relevant, and useful caricatures in that they describe central and dominating features in each discipline. Furthermore, by merging the words and premises of the different traditions (using them to "caricature" each other), a comparison of the different fields are not only possible, but also fair in the sense that both are given equally unfavorable treatment. Too many comparisons between these traditions usually favor one at the expense of the other, yielding a predictably skewed image. Lastly, the purpose of the caricature is neither the description of each discipline by itself nor the comparison between them, but to create a wide and abstract enough model of the foundation of grammars applicable to dialogism. Put metaphorically, this interdisciplinary analysis focuses on the omelet – not on the eggs broken nor on how they are scrambled.

For this project and similar endeavors en-route towards a dialogic grammar, the diatope data base is a good first step. But the diatope data base in itself is not a grammar. Apart from the brief discussion of grammar in section 2.1 and this analysis of language data bases, this article has not addressed the question "what can a dialogic grammar be?" Future studies need to address other fundamental questions about grammars in general and a dialogic grammar in particular *based on* an open account of their grammars' empirical foundation. Such studies will form a second step that presents "a system of rules [or

conventions] that in some explicit and well-defined way assign structural descriptions to” diatopes (cf. Chomsky 1965: 8). If such examples as the PageRank algorithm are truly indicative of the potential of a dialogic grammar, dialogism would do well to pursue such a second step and confront its resistance towards systematization.

References

- Bakhtin, M. 1981. Discourse in the novel. In M. Holquist (eds.) *The dialogic imagination* (pp. 259-422).
- Bakhtin, M. 1986. The problem of speech genres. In C. Emerson and M. Holquist (eds.) *Speech Genres and Other Late Essays* (pp. 60-102). Austin: University of Texas Press.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. 1969. Quine's empirical assumptions. In D. Davidson and J. Hintikka (eds.) *Words and objections. Essays on the work of W.V. Quine* (pp. 53-68). Dordrecht: Reidel.
- Evensen, L. 2002. Convention from Below: Negotiating Interaction and Culture in Argumentative Writing. *Written Communication* **19**(3), 382-413.
- Gasparov, B. 2008. Speech, Memory, and Meaning: Intertextuality in everyday language.
- Givon, T. 1979. *On understanding grammar*. New York: Academic Press.
- Holquist, M. 1990. *Dialogism: Bakhtin and His World*. London: Routledge.
- Hutchby, I., and Wooffitt, R. 2008. *Conversation Analysis*. Cambridge: Polity Press.
- Itkonen, E. 1996. Concerning the generative paradigm. *Journal of Pragmatics* **25**(4), 471-501.
- Linell, P. 1998. *Approaching dialogue: talk, interaction and contexts in dialogical perspectives*. Amsterdam: John Benjamins.
- Linell, P. 2005. Towards a dialogical linguistics. In *XII International Bakhtin Conference*, (pp. 157-172). Jyväskylä, Finland
- Nichols, J. 1984. Functional Theories of Grammar. *Annual Reviews in Anthropology* **13**(1), 97-117.
- Norén, N. 2007. Apokoinou in Swedish talk-in-interaction: A family of methods for grammatical construction and the resolving of local communicative projects. *Department of Culture and Communication: Linköping University*.
- Sacks, H., Schlegoff, E., and Jefferson, G. 1974. A Simplest Systematics for the Organisation of Turn-Taking in Conversation. *Language* **50**(4), 696-735.
- Wikipedia 2008. PageRank. Retrieved 19.09.2008 from <http://en.wikipedia.org/wiki/Pagerank>.

Ørstavik, I. 2008. Language shaping power: Bakhtin, Cassirer and Phenomenology. *XIIIth Bakhtin Conference*. London, ON.