



OPAALS PROJECT

Contract n° FP6-034824

WP6: Socio-Economic Constructivism and Language

**Del. 6.1 - Studies/papers on discourse organisation of
epistemic cultures – theoretical and methodological
analyses and practical interpolation**



Project funded by the European Community under
the “Information Society Technology” Programme

Contract Number: FP6-034824
Project Acronym: OPAALS

Deliverable N°: 6.1
Due date: May 2007
Delivery Date: August 2007

Short Description: <i>Deliverable D6.1 deals with the organisation of discourse in epistemic cultures and follows a twofold agenda: Firstly, its overall aim is to provide the necessary theoretical and methodological platform for further tasks carried out in work package 6 which are 'synthetically' connected in its linguistic focus and agenda. Secondly, the deliverable is structured as an integrative and comprehensive approach for OPAALS as it aims to develop an introductory framework regarding theories, methodologies, and practical interpolations for two overarching focal points of the project: (1)Communication processes -> discourse organisation; (2)Communities and knowledge -> epistemic cultures</i>
--

Author: Marco Bräuer, Andrea Crone, Oxana Lapteva, Frauke Zeller Partners contributed: Made available to: Consortium

Versioning		
Version	Date	Name, organisation
V1	07.08.2007	
V2	15.08.2007	

Quality check Internal Reviewers: Andrea Nicolai (T6-Eco), Josef Wallmannsberger (UniKassel)

Dependences:

Work Packages	4, 7, 8, 10
OPAALS Partners	LSE, CreateNet, T6-Eco, WIT, UL
Domains	Social science: Linguistics, computational linguistics, organisational communication, empirical media studies
Targets	All partners



This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 2.5 License. To view a copy of this license, visit : <http://creativecommons.org/licenses/by-nc-sa/2.5/> or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

Table of Contents

Index of Tables.....	4
Index of Figures.....	4
0) Introduction.....	5
0.1) Overall Introduction and Interpolations.....	5
0.2) Deliverable 6.1 Structure.....	7
1) Discourse Organisation	10
1.1) Introduction.....	10
1.2) Discourse Models and Paradigms.....	12
1.2.1) Discourse(s) in Computational Linguistics.....	12
1.2.2) Philosophy of Language.....	14
1.2.3) Social Sciences.....	15
1.2.4) Linguistics.....	16
1.3) Domains of Applied Discourse Studies.....	18
1.3.1) Organisational Discourse	18
1.3.2) Information and Communication Technology & Computer-Mediated Communication	20
1.3.3) Formal and semi-formal languages.....	22
1.3.4) Domain-Specific Languages and Languages for Special Purposes.....	24
1.4) Summary.....	31
2) Epistemic Cultures.....	32
2.1) OPAALS and the notion of epistemic cultures.....	32
2.1.1) Interdisciplinarity.....	34
2.1.2) Open Knowledge Community.....	35
2.1.3) Autopoiesis.....	35
2.1.4) Reflectivity.....	36
2.2) Knowledge cultures.....	36
2.2.1) Epistemic cultures.....	36
2.2.2) Epistemic communities and Community of Practice	38
2.2.3) Organisational Communication.....	39
2.2.4) Genre Concept and Media richness.....	44
2.3) Summary.....	51
3) Paradigms of epistemic cultures	52
3.1) Introduction.....	52
3.1.1) Paradigms.....	54
3.1.2) Epistemological assumptions	54
3.1.3) Research Design.....	57
3.1.4) Research methods	59
3.2) Why to Combine? - Theoretical and Methodological Implications.....	61
3.3) Pragmatism.....	62
3.3.1) Rational for mixed method research	63
3.3.2) Mixed method research design.....	64
3.3.3) Philosophical debates.....	67
3.4) Summary.....	70
4) Sociolinguistic Framework.....	72
4.1) Introduction	72
4.2) Framework Domains: Multidimensionality of Analysis.....	73
4.2.1) Language.....	73
4.2.2) Knowledge	73
4.2.3) Communication.....	75

4.2.4) Network.....	76
5) Conclusion.....	77
6) Bibliography.....	79

Index of Tables

Table 1. Formalism and functionalism view of discourse and language.....	11
Table 2. Formal and semi-formal domains.....	23
Table 3. Discourse Community and Community of Practice.....	29
Table 4. Qualitative and Quantitative Paradigm Assumptions.....	56
Table 5. Paradigm of pragmatism.....	62

Index of Figures

Figure 1. Work package 6 dependencies in Phase I.....	7
Figure 2. Factors influencing discourse within the OPAALS scientific community.....	25
Figure 3. Process of deduction in linear quantitative research design.....	58
Figure 4. Circular strategy of qualitative inductive research design.....	59
Figure 5. Methodological research design in the pragmatic paradigm.....	65
Figure 6. Mixed research process model.....	68
Figure 7. Sociolinguistic Framework.....	73
Figure 8. OPAALS' living organism.....	76

0) Introduction

0.1) Overall Introduction and Interpolations

Deliverable D6.1 deals with the organisation of discourse in epistemic cultures and follows a twofold agenda: Firstly, its overall aim is to provide the necessary theoretical and methodological platform for further tasks carried out in work package 6 which are 'synthetically' connected in its linguistic focus and agenda. Secondly, the deliverable is structured as an integrative and comprehensive approach for OPAALS as it aims to develop an introductory framework regarding theories, methodologies, and practical interpolations for two overarching focal points of the project:

- (1) Communication processes -> *discourse organisation*
- (2) Communities and knowledge -> *epistemic cultures*

OPAALS' first objective (among 8, see Description of Work) is to develop an integrated theoretical foundation for digital ecosystems research, for which this deliverable provides a first social science perspective. The OPAALS Network of Excellence also heads towards the development of a sustainable interdisciplinary research community which focuses on knowledge production and management. One claim is that in order to achieve sustainable digital business ecosystems, the collaborative processes and ICTs that underpin the continuous creation, formalisation, and sharing of knowledge need to be understood in depth. This claim is integrated by the two aforementioned focal points, as 'communities' represent the framework in which collaborative processes can be found, and communication processes the “socio-constructivist” part of the framework.

Approaching communication processes and discourse as a socio-constructivist medium implies that these constitute the main vehicle for information processing and contextualisation which is the cornerstone of knowledge creation. Therefore, it is necessary to analyse existing communication processes within the multiple communities of OPAALS (i.e. research teams, domains, etc.), how they 'organise' collaboration and apply different discourse patterns. The notion of 'discourse' organisation integrates these points and also refers to the notion of organisations as “discursive” units, i.e. they are shaped and perceived by means of different forms of discourse or communication in general. The notion of organisations or communities as discourse units also integrates the two distinct genres of language: natural and formal languages. The linguistic agenda of work package 6 allows to expand the traditional approach to communication studies in organisations by means of the computational linguistics domain: It addresses the interplay of natural and formal languages, that is to say combines natural language discourse units with processable and computable formal language representations. This complementary approach does not only play an important role in work package 6 but also influences work package 2 (“Automatic code generation from models”), work package 4 (“Distributed accountability, identity, and trust”), and work package 7 (“Community networks and digital ecosystems”). Work package 2 integrates a predominantly

formal language focus with natural language as its 'laboratory object of analysis' and aims to provide a unified theory for the automatic creation of programmes from a specification written in human natural language. To achieve this, work package 2 systematically applies computer science approaches and theories to natural sciences combined with social science approaches from epistemology, epistemic communities, and computational linguistics. The main objective of work package 4 is to define models for distributed identity, accountability, and trust for the digital ecosystem that facilitates behavioural and economic collaborations. Whereas its overall expertise lies in computer science, it clearly encompasses theories from the social science domain regarding the definition and design of its models. These benefit from a comprehensive notion of communities regarding important 'social' factors such as trust and identity. Work package 7, finally, approaches the field of community networks and, among other topics, information infrastructures by means of a socio-technical framework which integrates both social and technological factors. Methodologies from computational linguistics and computer-mediated communication offer therefore additional approaches that focus on communication structures in socio-technically driven community networks.

'Communities and knowledge' represents the second focal point and is addressed by means of an interdisciplinary elaboration of 'epistemic cultures'. Work package 6's language agenda aims at a systematic approach for the representation of interdisciplinary communities and research networks representing aggregated systems of knowledge production. Hence, the framework of communities and community building must be understood regarding the development of networks, their processes and socio-cultural implications. The notion of epistemic cultures allows to apply a broad research spectrum on communities and knowledge, which means that it does not only focus on strict sociological methodologies and theories. Epistemic cultures also leads to the field of domain specific languages, which integrates questions regarding policies, identity, trust, and governance. Therefore, there are additional connections to work package 7 ("Community networks and digital ecosystems"), work package 8 ("Open Source and Open Knowledge"), and work package 10 ("Sustainable community building").

From a meta-perspective, work package 6 covers its language agenda from three complementary points of view:

1) Language as subject of analysis

- Technical languages ('agorising' metaphors); discourse organisation (T6.1, T6.2);
- Symbolic ecosystems (T6.3)

2) Language as a formalised tool for construction

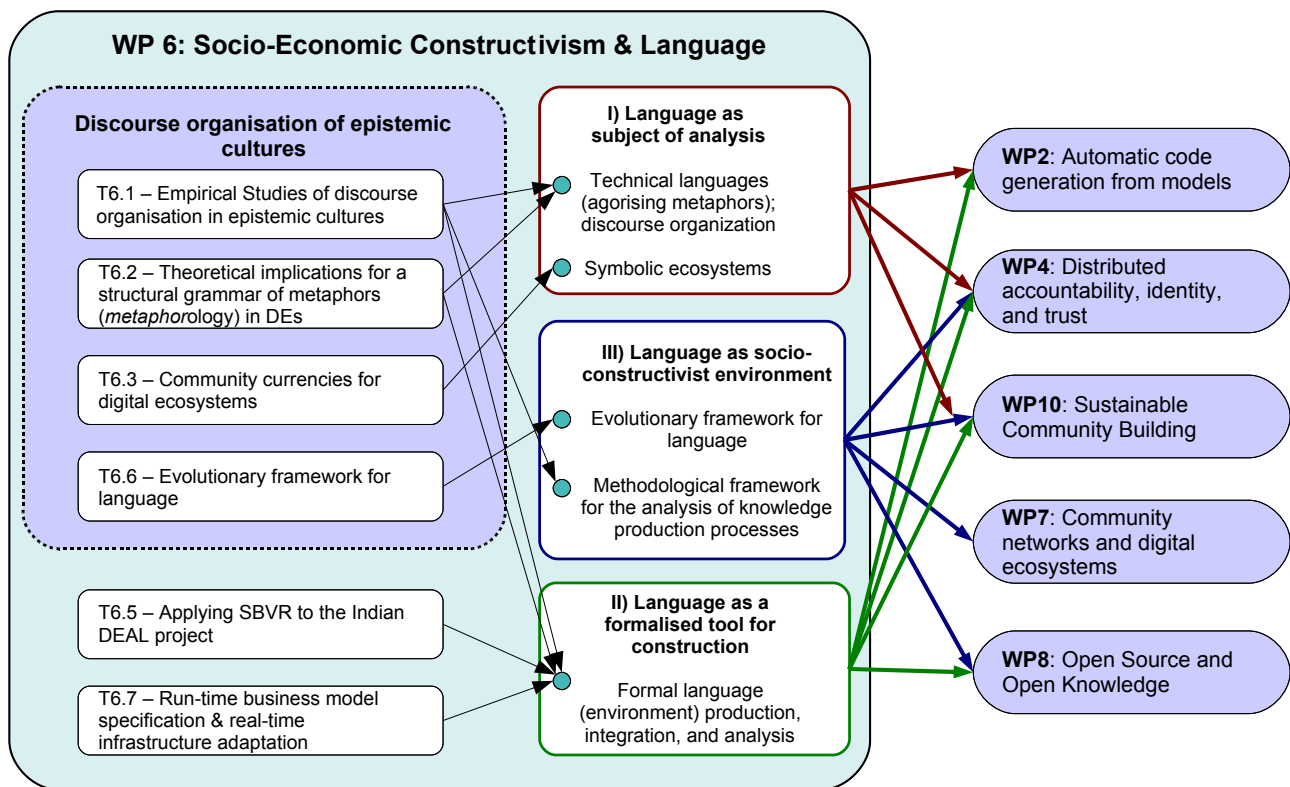
- Formal language (environment) production, integration, and analysis (T6.1, T6.5, T6.7)

3) Language as socio-constructivist environment

- Evolutionary framework for language (T6.6)
- Methodological framework for the analysis of knowledge production processes (T6.1, T6.4)

The following figure (Figure 1) shows the content related dependencies between the different tasks in work package 6 of Phase I and connections to aforementioned work packages, which result in a holistic and extended framework of applied language studies in Digital Ecosystems.

Figure 1. Work package 6 dependencies in Phase I



0.2) Deliverable 6.1 Structure

The systematic approach of work package 6 and particularly this deliverable implies sound and exhaustive research regarding state-of-the-art approaches and studies as a first step. The consecutive step includes the realisation of further studies and publications of the results as well as research outcomes which use this deliverable as point of departure.

Deliverable 6.1 provides a summary of the outcomes of the first step as it encompasses the analysis of a rich set of studies, research outcomes, etc. in those fields which on the one hand represent important approaches for social science research undertakings within OPAALS and, on the other hand, connection points for OPAALS' computer science and natural science domain. It therefore can be regarded as a research and work matrix for the second step – the realisation of further studies and publications of the results – while the second step itself is implemented in different tasks of work package 6 and further work packages that address language questions.

Regarding the two focal points ('communication processes', 'communities'), different aspects of 'communication' and 'discourse' are addressed in this deliverable in chapter 1 by means of an introductory discussion of the terminology. Taking into account the main focus of this deliverable which addresses communication processes as 'discourse' units, the deliverable offers the necessary theoretical underpinnings of discourse studies, followed by selected studies and conclusive discussions of different fields regarding discourse and language (e.g. ICTs and computer-mediated communication, formal and semi-formal languages, domain specific languages).

Chapter 2 provides an introduction to the field of epistemic cultures and its connection and implications to/with OPAALS. Using the notion of 'epistemology' and/or 'knowledge' as “redeeming” feature of both this deliverable and OPAALS, different approaches to communities and knowledge are also discussed in this chapter drawing on surveys conducted in the fields of, for example, communities of practice or organisational communication. The chapter concludes with an extensive discussion on 'genre' and 'genre systems' which are used in multiple research designs in order to integrate a connectionist view on organisations, communication, and communicative tools. These studies offer a broad range of insights for all OPAALS partners regarding possible ways of structuring their own research or – on a meta-level – taking a systematic approach to any tasks relating to community, collaboration, technology, and language.

Chapter 3 focuses on research designs and methodologies in the field of epistemic cultures and offers a rich set of possible connection points as well as practical points of departure for further tasks and work packages. It therefore represents an introductory tool-box from social science for all OPAALS partners for practical interdisciplinary work. Chapter 3 also offers a detailed view on the studies depicted in chapters 1 and 2 regarding their applied methods and therefore aims at making social science approaches more visible and understandable to non-social scientists.

The last main chapter (4) discusses an approach which is being used in work package 6, the “Sociolinguistic Framework”. This approach is currently applied for empirical research in OPAALS focusing on communities, knowledge, and language/communication. It is therefore closely connected with tasks 6.1 and 6.2 in work package 6, and also with several tasks in work package 10 which deal with community building in a socio-technical framework (tasks 10.6 and 10.7).

To sum up, regarding the aforementioned goals, objectives and focal points of work package 6 and other work packages in this Network of Excellence, deliverable 6.1 aims:

- To address the two overarching concepts of OPAALS: language and community.
- To foster the understanding of a social science framework within the OPAALS community.
- To initiate interdisciplinary discourses.

1) Discourse Organisation

1.1) Introduction

The concept of discourse is generally defined in two ways: discourse as a particular unit of language (above the sentence) or discourse as a particular focus on language use (Schiffrin, 1994). These definitions of discourse reflect two different paradigms: functionalism and formalism. The major difference between these two approaches in the context of discourse is the definition of language. The formalists (for example, Chomsky) define it as a mental phenomenon, whereas the functionalists (for example, Halliday) tend to see it primarily as a social phenomenon.

Further characteristics and differences of formalist and functionalist approaches in regard to a definition and theory of discourse are shown in Table 1.

Table 1. Formalism and functionalism view of discourse and language

	Formalism	Functionalism
Linguistic universals	Deriving from a common genetic linguistic inheritance of the human species.	Deriving from the universality of the uses to which language is put in human society.
Acquisition of language	A build-in-human capacity to learn language.	Development of the humans' communicative needs and abilities in society.
Language	Studying language as an autonomous system.	Studying language in relation to its social function.
Discourse	A level of structure higher than the sentence, or higher than another unit of text.	Interdependent from social life, such that its analysis necessarily intersects with meanings, activities, and systems outside of itself.

Functionalists view discourse as a primary feature of social reality. Through understanding discourse we are able to understand our reality and ourselves (Phillips & Hardy, 2002). Phillips and Hardy (2002) point out that researchers and the research community are parts of the constructive effects of discourse by using language, producing text, and drawing on discourses. Some authors define discourse as the actual practices of talking and writing (Woodilla, 1998), others use the term

in a more specific way. Parker defines discourse as an interrelated set of texts including their production, dissemination, and reception (Parker, 1992). Chalaby (1996) identifies the concept of discourse as a social reality that cannot be explored and understood without the discourse. Phillips and Hardy (2002) have introduced the multidimensionality of discourse in the sense that it connects texts to discourse, locating them in a historical and social context by which they refer to the particular actors, relationships, and practices that characterise the situation under study. An important point from Phillips and Hardy is that discourse analysis examines how language constructs phenomena, not how it reflects and reveals it (Phillips & Hardy, 2002). In the framework of language and discourse, Georgakopoulou and Goutsos (1997) focus on communication. Language usage in communicational processes is not exclusively produced through individual units (sounds, words, or sentences), but by combination of these language units. The authors found systematic co-patternings between form, content, function and context of people's discourse activities.

A standard term that captures these co-patternings is "genre", which is also well known outside linguistics. Genre is often defined as a classificatory concept referring to a class of communicative events, that are characterised by a variety of shared formal, functional and contextual properties (Georgakopoulou & Goutsos, 1997). The term genre is not only determined by reference to form such as lexical or grammatical patterns and content; it is also determined by reference to social and cognitive criteria, such as norms, conventions, rules of use, schemata and, on the whole, our perceptions of and expectations about textual boundaries (Georgakopoulou & Goutsos, 1997). From this point of view, each society can have its own particular configuration of genres, often called "genre systems", which are in particular relation to each other. The studies of genre and genre systems see texts as socially produced and patterned (Bamford & Bondi, 2005). These studies focus on different genres and try to:

- analyse textual patterns and features which characterise and define the genre;
- analyse how texts of a certain genre are socially constructed and situated;
- present how texts relate to the society in which they are produced and consumed;
- identify genre systems existing within a particular community.

In the case of academic discourse examples for text genres or genre systems could be books, research articles, working or conference papers, lectures, research proposals, seminars, workshops, reports, etc. According to Bamford and Bondi (2005), genres may no longer be separable from one another as was once supposed: A research article for example may be the end result of a series of other genres - a conference paper, a seminar, conversations with colleagues, research reports, etc. Moreover, scientists may produce newspaper articles, television interviews and other hybrid genres which are a combination of professional and non-professional genres (Bamford & Bondi, 2005).

Genres are defined as dynamic systems because “they are embedded in the discursive practices of disciplinary cultures which in turn are part of a society and culture which is in a state of constant fix” (Bamford & Bondi, 2005, p. IX).

In conjunction with on-line discourse communities and especially OPAALS' as an interdisciplinary and dispersed organisation, genre systems can be seen as interactive, dynamic and distributed over several groups, communities or cultures. The study of the different genre systems used by the multiple teams and actors within OPAALS would reveal (a) communicative patterns connected to certain genres; and (b) expectations and pre-dispositions of and towards the usage of certain genre systems (i.e. reputation and acknowledgement of an article published in a weblog vs. in a peer-reviewed international journal). The results could be used in order to develop the Open Knowledge Space, combining the different preferences and domain 'policies' regarding text/discourse production. Furthermore, the notion of genres as dynamic systems is especially important in on-line communities, where additional expectations might arise, based on ICT which enables cross-media publishing, that cannot be met due to the intrinsic different attitude towards the different genre systems. For example, the expectation from a content management system to feed in automatically content from a personal blog to the community news forum: A news forum is often perceived as not including personal information but to rather inform about general facts such as meetings, events, etc.

1.2) Discourse Models and Paradigms

This sub-chapter provides a brief overview on discourse models and paradigms within different disciplines. It therefore stresses the interdisciplinary nature of discourse and its multiple application domains, which can contribute to OPAALS varying foci and tasks. Sub-chapter 1.3 then discusses a selection of concrete domains where discourse studies are applied.

1.2.1) Discourse(s) in Computational Linguistics

Computational linguistics is a domain, which covers a vast range of approaches and intellectual endeavours. Kay (2004) sets its origin or rather point of departure with Warren Weaver's memorandum in 1949 where he suggested that translation by machine might be possible. But it was not until the 1960s when the name computational linguistics first appeared in the title of the journal *Mechanical Translation and Computational Linguistics*. Therefore, a strong set of methodologies comes from the engineering domain, which deals with topics such as machine translation or speech processing (see Mitkov, 2004; Jurafsky, 2000). Digital humanities are also inspired by computational linguistics and depict computational approaches to language processing and

interpretation which are inspired by methodologies and paradigms from the Arts or Humanities, which originate from machine supported corpus studies that were inspired by the analytic and indexing undertaking of Father Roberto Busa (Schreibman, Siemens & Unsworth, 2004).

Regarding the field of discourse analysis and the application of discourse models, computational linguistics focuses on models of discourse in order to construct computational systems capable of natural language processing (i.e. interpreting and generating). Webber (2003) defines two general aims of computational work on discourse and dialogue:

- [T]hat of modeling human understanding and generation of natural language in terms of a system of computational processes. Work in this area is usually called computational linguistics.
- that of enabling computers to analyze and generate natural language in order to provide a useful service. Work in this area has been called applied natural language processing, natural language engineering, or more recently language technology. (p. 798)

A structuralist approach to discourse can be depicted by a computational linguist's conviction that any coherent discourse has to contain structure, and that recognizing the structure is a crucial component of comprehending discourse (Moore & Wiemer-Hastings, 2003). Three types of structure can be differentiated: Intentional structure, informational structure, and attentional structure. Intentional structure deals with the roles utterances play in the speaker's communicative plan to achieve desired effects regarding the hearer. Grosz and Sidner (1986) pay particular attention to the intentional structure of utterances as they consider them as an important source for coherence. They also claim that "the structure of any discourse is a composite of three distinct but inter-acting components: The structure of the actual sequence of utterances in the discourse; a structure of intentions; an attentional state" (p. 176). It has to be stressed, though, that the authors depict intention and attention as non-linguistic notions. They also point out that their concern is not discourse meaning but only discourse structure, hence differentiate between meaning and structure: "An adequate theory of discourse meaning needs to rest at least partially on an adequate theory of discourse structure" (ibid).

Kamp addresses directly the problem of meaning processing regarding dynamic semantic and Discourse Representation Theory. Main concern of DR theories is the context dependence of meaning and the reciprocal interaction between context and utterance and its central concept information. DR theories also represent – according to Kamp – a paradigm shift in formal semantics, from its traditional focus on reference and truth to the central concept of information. Hence, the meaning of a sentence is not its truth conditions but its information change potential, leveraged by modifying given contexts or information states (Kamp).

Wahlster (1991; Kobsa & Wahlster 1988) has a different focal point and addresses the social and/or individual component of meaning in a computational model of discourse: He uses a discriminative approach and differentiates between user models and discourse models. A user model is depicted as a knowledge source that stores all explicit assumptions about the user that may be relevant to the overall dialogue behaviour of the system. Its main functions are, besides the incremental establishment of the model, to maintain consistency and feed other components of the system with user information. He defines discourse models as a knowledge source with the system's representation of the syntax, semantics, and pragmatics of a proceeding dialogue (Wahlster, 1991). The important accompanying assumption in Wahlster's models is the notion of dynamics and evolvment: The user model is specifically bound to a constant revision and updating process, where entries (such as beliefs) are updated or even deleted. Whereas the discourse model lacks the "notion of logical consistency that is important for belief revision", it is also highly dynamic as it developed "like an annotated trace of the various levels of the system's processing involved in understanding the user's utterances and generating its own dialogue contributions" (p. 47).

The idea of including computational transcripts of individual beliefs (and desires, and intentions) also connects to the notion of meaning and how it is built in/through language. What is easily being overlooked is the danger of mixing up two intrinsically different discussions on meaning: The individual's sense-making-processes and the social construction of meaning. The latter refers to the structuralist tradition and de Saussure's definition of *langage*, specifically *langue*. *Langue* depicts the language system at a given time in a given community (synchronic approach vs. diachronic studies of language) apart from the individual's manifestations, which is called *parole*. The language system consists of signs that are depicted as a dualist, arbitrary, and static relation between signifier and signified: The signifier is a sound image of a certain sign (e.g. word) and the signified is the concept associated with the sound image (de Saussure, 2003). This means that the "meaning" of the signifier is socially agreed upon and constructed in a given community.

1.2.2) Philosophy of Language

The notion of meaning and discourse models also refers to the field of philosophy of language, which offers a vast set of ideas and models on discourse and language. Frege, for example, discussed "meaning" in more detail. He argued that in order to understand how language can represent reality we require a distinction between *Sinn* and *Bedeutung* (sense and reference) instead of relying solely on our intuitive notion of meaning (Miller, 2003). Frege defines: "The *Bedeutung* or semantic value [reference] of any expression is that feature of it which determines whether sentences in which it occurs are true or false" (cited in Miller, 2003, p. 265).

In his original text *Über Sinn und Bedeutung*, Frege (1892) expresses the crux of the dilemma of any "meaningful" undertaking into the computational processing of natural language:

Der Sinn eines Eigennames wird von jedem erfaßt, der die Sprache oder das Ganze von Bezeichnungen hinreichend kennt, der er angehört; damit ist die Bedeutung aber, falls sie vorhanden ist, doch immer nur einseitig beleuchtet. Zu einer allseitigen Erkenntnis der Bedeutung würde gehören, daß wir von jedem gegebenen Sinne sogleich angeben könnten, ob er zu ihr gehöre. Dahin gelangen wir nie. (Frege, 1892, p. 27)

The translation reads:

The sense of a proper name is grasped by everybody who is sufficiently familiar with the language or totality of designations to which it belongs; but this serves to illuminate only a single aspect of the reference, supposing it to have one. Comprehensive knowledge of the reference would require us to be able to say immediately whether any given sense belongs to it. To such knowledge we never attain. (Frege, 1960, p. 58)

By using the well-known example of morning-star and evening-star, Frege claims that both “signs” convey the same reference, but not the same sense. Hence, one can designate a specific sense to each sign and to that sense a corresponding reference. But there is no one-to-one relation between reference and sign: there are different signs that can belong to one reference. According to this definition, reference seems to be the “invariant” constituent whereas sense is open to change. However, this would overlook the fact that natural language is a highly dynamic system, and there can well be reference changes. For example, the sign “mouse” conveyed only one reference until Douglas Engelbart invented a pointing device for the interaction man/woman with the computer in 1964.

1.2.3) Social Sciences

Foucault’s work on discourse in “The Archaeology of Knowledge” (1972) is regarded as highly influential in social science. He uses the term discourse (abstractly) for a “domain of statements”, as a “count” noun, for groups of statements or for the rules which govern such a group of statements. Fairclough (2003) summarizes:

The analysis of discourse for Foucault is the analysis of the domain of ‘state-ments’ – that is, of texts, and of utterances as constituent elements of texts. But that does not mean a concern with detailed analysis of texts – the concern is more a matter of discerning the rules which ‘govern’ bodies of texts and utterances. (p. 123)

Instead of “drawing up tables of differences (as linguists do)”, Foucault (1972) uses the term systems of dispersion for the analysis of discourse(s):

Whenever one can describe, between a number of statements, such a system of dispersion, whenever, between objects, types of statement, concepts or thematic choices, one can define a regularity [...], we will say, for the sake of convenience, that we are dealing with a discursive formation [...]. The conditions to which the elements of this division (objects, mode of statement, concepts, thematic choices) are subjected we shall call the rules of formation. (p. 38)

He differentiates four discursive formations: formation of objects, formation of enunciative modalities, formation of concepts, and formation of strategies. He explains discourse mainly by means of discursive mechanisms where power relations play a vital role. Discourse uses a notion of the “openness”, “external” as its constitutive element, because discourses *ad definitionem* do not consist of a singular text but rather of an ensemble of texts. Only the relations of those texts to each other constitute discourse: these could be relations of power, competition, opposition or even coalition.

Chalaby (1996) suggests a shift of perspective regarding a social science approach to discourse studies: Instead of using a linguistic motivated notion of discourse as linguistic and textual phenomenon, the (social) production conditions of discourse need to be addressed in the first place:

[...] a discourse is defined by the external – external because social – conditions of formation and production that make a class of texts a coherent and constructed sociological object [...] texts belong to the same discourse as long as particular relations of production condition the set of discursive practices which produces a class of texts [...]. (Chalaby, 1996, p. 690)

1.2.4) Linguistics

The term discourse is used in a range of different fields with different meanings. It is claimed that linguists use the somewhat broadest definition of discourse by defining it as anything “beyond the sentence”. Schiffrin, Tannen and Hamilton (2001) summarise the surplus of definitions by using three main categories: “(1) anything beyond the sentence, (2) language use, and (3) a broader range of social practice that includes non-linguistic and non-specific instances of language” (p. 1).

To summarise different views on discourse, Schiffrin (1994) identifies six different approaches to discourse analysis:

1. **Speech Act Theory:** The focus of Speech Act Theory lies primarily on utterances and conditions underlying their production and understanding in the context of different communicational situations and actions. The major concerns of this theory are the language functions and how meaning of what we are saying and our actions are related to it. It emphasises the importance of the differentiation between different circumstances under which we perform our actions and knowledge of the existing conditions and rules.

2. Interactional sociolinguistics: Interactional sociolinguistics is based on different disciplines such as anthropology, sociology, and linguistics. This approach addresses a variety of factors related to culture, society, and language. The major aspects of research within the field of interactional sociolinguistics are the way people share grammatical knowledge of language and how context influences the production and understanding of what is said.
3. Ethnography of communication: This approach relies on studies and theories in anthropology and linguistics. These two disciplines are often seen as clearly distinct from one another in regard of their data, problems, methods, and theories. Their crosspoint of interest, however, is the concept of communication. Generally, an ethnographic approach to discourse seeks to discover and analyse the structures and functions of communication and communicative actions that provide the “space” for the language use in different speech situations, events, and acts.
4. Pragmatics: Pragmatics is another approach to discourse and deals with three different concepts – meaning, context, and communication. A central aspect of pragmatics is context and how it influences the way we interpret texts and/or discourse units.
5. Conversation Analysis: Major factors of conversation analysis are social order and two dimensions of language (how it creates context and it is created by social context). This approach is close to the ethnography of communication approach in its “concern with human knowledge and its belief that no detail of conversation (or interaction) can be neglected a priori as unimportant” (Schiffrin, 1994, p. 232).
6. Variation Analysis: The roots of variation analysis lie in quantitative studies of linguistic change and variation. Studies using this approach primarily focus on semantically equivalent variants and try to discover their social and linguistic patterns. Although variation analysis includes social aspects of the text and its context, it primarily focuses on the specific linguistic features.

A different approach to linguistic discourse analysis can be found in semiotic engineering and/or computational semiotics, which deals with “the principles, materials, processes, effects, and the possibilities for producing meaningful interactive computer system discourse” (de Souza, 2005, p. 4). Conversely to traditional user-centred design studies, de Souza applies the same role to both users and designers of software systems: they are interlocutors in an overall communicative process. This process ascribes to the designers the communicative sender role, where they have to tell users the specific meaning they seek by the artefact they have created, and users must participate in the communication process by understanding/responding to what they had been told. Hence, semiotic engineering can be depicted as a highly reflective theory; the designer is integrated into human-computer-interaction processes and assigned to a role as important as that of the user’s (de Souza, 2005).

A different computational semiotics approach is the concept of the algorithmic sign. Andersen (2003) regards computers as a natural subject of semiotics, as they function as media that process representations. He and Nake use the term algorithmic sign to denote the double nature of the computer-based sign:

[O]n the one hand a normal sign based on intentionality, on the other hand a mechanical process relying on causes and effects. The algorithmic sign – or the computer artefact viewed as a sign – appears as a phenomenon of a new kind: it is a description that can perform its own contents. The algorithmic sign shares with all other signs the property of being something that stands for something else and being interpreted by humans. But in addition the algorithmic sign can become active, it can be set in motion, can run, be executed. (Andersen, 2003, p. 173)

Nake (cited in Andersen, 2003) claims that HCI is determined by two concurrent kinds of processes, which are firstly the signal processes inside the machine (based on cause and effect) and secondly the sign processes among the users (based on free interaction). This approach clearly promotes the computer as a mere tool to a significant sign system process itself. It means that we can find a technological dimension of sign processes (or discourse) as a concrete and potential agent (actor) integrated.

1.3) Domains of Applied Discourse Studies

1.3.1) Organisational Discourse

In communication processes we give meaning to our experience and express it through the medium we chose to use (phone, email, face-to-face, or others). In our environment, in collaboration and communication with others, we often use language as a tool to describe and represent reality. Thus, the ability to understand the language of others depends not only on the vocabulary and content, it is „also context and a way to re-contextualise content“ (Boje et al., 2004, p. 571). Additionally, language does not only have the function of reporting and describing reality. Aspects in creation and creativity of language production and usage provide an additional dimension of its complexity: “... what we create in language “uses us” in that it provides a point of view (a context) within which we “know” reality and orient our actions” (p. 571).

One way to contextualise discourse is to link it to certain community forms, such as organisations. Boje et al. (2004) introduce organisations as “phenomena in and of language” and as collaborative and contending discourses. The authors argue that there are several ways or forms of engagement within organisations from a linguistic and discursive point of view: foci, methods, levels, and modes. The first form of engagement (*focus*) regards the use of language as a “vehicle for analysing and exploring organisations and organising (language as a means to an end) or treating

organisations as sites for language analysis (language as an end in itself)” (Boje et al., 2004, p. 571). In the context of the *methods* of engagement, there is a variety of methodological directions such as conversation analysis, content analysis, critical discourse analysis, ethnomethodology, etc. Regarding *levels* of engagement the authors differentiate between three levels of analysis: micro (e.g. discrete organisational episodes or conversations), meso (e.g., broader patterns and networks of organisational interaction) and macro (e.g., grand narratives and metadiscourses with wider social implications). And finally, in the context of *modes* of engagement, the researchers can “interrogate organisations and organising processes by privileging monologic, dialogic, or polyphonic perspectives” (Boje et al., 2004, p. 571). An important aspect of this study is that discourse is viewed as “the intermingled play of differences in meanings mediated through socially constructed language practices (some of which are hegemonic), especially in genres of verbal utterances such as stories and conversations, as well as in material inscriptions in other texts” (Boje et al., 2004, p. 572).

Putnam and Fairhurst (2001) focus on language analysis and discourse in organisations and define discourse:

[...] as a way of knowing or a perspective for understanding organizational life. It is a lens or a point of entry for seeing, learning, and understanding ongoing events. As a lens, it provides a unique way to focus on the subtle aspects of organizing and to determine what is figure and ground in the framing of organizational events. (p. 79)

Weick (1979) uses a more abstract advance and depicts communication as the core process of organising and focuses on the organisation of discourse models. In his book “Sensemaking in Organizations” (1995), Weick defines organisations as social contexts for sense-making which points to discussions of environmental aspects and boundaries of organisations: “Given bounded rationality and environmental complexity, sensemaking tends to crystallize into cognitive frames that reduce ambiguity and facilitate decision making” (Santos & Eisenhardt, 2005, p. 500). Other advances in organisational studies directly address the lack of traditional boundaries and organisational environments and approach this “shortcoming” by discussing phenomena of locally dispersed organisations. They are redefined and relabelled as “hypermedia networks” or “knowledge networks” (Brown & Duguid, 1991; Howard, 2002; Podolny & Page, 1998), and connect to the interdisciplinary fields of sociology of technology, social software, or social networks.

The concept of organisations as a locally dispersed being-in-the-world depicts the challenges of any “hands-on”-approach regarding the analysis of discourse patterns and structures taking into account the lack of tangible frames or demarcation lines. For those organisations, language (as a more general notion of discourse) can act in three ways as a possible bond to fill the spatial absence:

1. Language as a means to express and create a certain notion of space (e.g. deixis, meta-phors) and therefore to bootstrap group work within networked organisations.
2. Language regarding the practical interface design of technological applications, that is to say a means to enhance and facilitate human-computer-interaction.
3. Language as a precipitate, which can be (electronically) studied in order to analyse aspects of work flow, inter-group relations, and management questions of organisations .

These three core roles define language as a tool: (a) and (b) as a tool for creation, and (c) as a descriptive tool or tool for reflection. In the case of the OPAALS focal points and objectives, however, making solely use of the language-as-a-tool metaphor would be by no means sufficient. In order to build a sustainable community for digital ecosystems, an analysis of how language is used reflects only the first step and needs to be combined with a broader socio-technical framework that takes into account the development of on-line spaces and software for knowledge generation and management. The different approaches to discourse and their models reflects an important part of the development process, where discourse is being represented in its contextual (socio-cultural) as well as instrumental dimension. For example, the fields of computational linguistics and philosophy of language can contribute significantly to formal language modelling environments, that are addressed in work package 2 and task 6.6 in work package 6.

1.3.2) Information and Communication Technology & Computer-Mediated Communication

In the course of the past years, Information and Communication Technology (ICT) has not only changed science and work life but also the research focus on text and discourse. ICT facilitates collaboration in dispersed research teams, provides new publication platforms (open access, online-journals, web logs, and databases) and enriches interpersonal communication with new tools. Moreover, ICT is regarded as a suitable means for community building, since through implementation and usage of different tools existing within ICT, community members are able to access information and knowledge. In this context, a new field study – computer-mediated communication (CMC) – gained increasingly interest in the research areas of epistemic cultures, discourse studies, linguistics, communication studies, etc .

CMC provides the ability to exchange different information (data) in networked environments. Typical computer-mediated formats used in organisations and communities are instant messages, e-mails, chat rooms, forums, etc. Generally, research on CMC in communities and epistemic cultures focuses on the social effects of different computer-supported communication technologies. Additionally, language aspects in computer-mediated context represent an important field of research. The study of language in these contexts is typically based on text-based forms of CMC, and is often referred to as "computer-mediated discourse analysis".

One of the tendencies in CMC related to language and discourse described by Herring (2001) is *computer-mediated discourse* (CMD) (a short historical sketch on the development of CMD can be found in Herring, 2003). CMD is “the communication produced when human beings interact with one another by transmitting messages via networked computers” (Herring, 2001, p. 612). As a part of CMC, the CMD approach focuses on language and language use in computer networked environments. A broad range of computer-mediated communication tools are text-based and take a variety of forms, such as e-mail, chat, forum, etc. Their linguistic properties vary “depending on the kind of messaging system used and the social and cultural context” (p. 612). Herring disagrees with a popular perception that computer-mediated language is less correct, complex, and coherent than standard written language. Despite the fact that computer-mediated language often contains non-standard features, “only a relatively small percentage of such features appears to be errors caused by inattention or lack of knowledge of the standard language forms” (p. 617). Often the users economise on typing effort, or mimic features existing in face-to-face communication, or simply show their creativity. CMD is also based on methodologies from critical discourse analysis (CDA) which takes as point of departure the notion of discourse use relative to social, political and cultural formations, and discusses interrelations between language use and power or ideology (see for example Fairclough, 1995; Fairclough, 2001; van Dijk, 1993; van Dijk, 2000; Wodak & Meyer, 2001; Bourdieu, 2006). But what is also relevant for our agenda is the rise of importance of discourse in the field of epistemology – in the theorising of knowledge: “The question of how we build knowledge has come more to the fore, and this is where issues to do with language and linguistic representation come into focus” (Jaworski & Coupland, 2006, p. 3). This describes a more formalised approach to discourse, which does not neglect the social dimension but includes an analytic approach towards attempts of formal classifications and relationships of knowledge which are expressed through language.

Language usage in computer-mediated environments such as OPAALS is highly variable: social factors, for example, participants' demographics and situational contexts influence the linguistic choices of the user and are therefore important in the different areas of research existing in the OPAALS project. Tasks 6.1 and 6.2 take CMC and CMD approaches into account in included them in their empirical survey regarding the different epistemic cultures within the Network of Excellence. Particularly the notion of an evolving community can be mirrored by changing communicative behaviour, such as the broadening of communication partners beyond the personal research team, or the usage of new communication tools. Therefore, the notion of the evolving community will be analysed by means of a long-term study, which also aims at revealing changing associations regarding important concepts and common goals within the OPAALS community.

1.3.3) Formal and semi-formal languages

One of the major problems related to ICT and software design is the development of proper and useful systems that can be used as an “interpreter” between the human and the machine. Considering discourse organisations in epistemic cultures, the modelling of knowledge and its implementation in software systems is likewise difficult and often not sufficiently supported by a proper socio-technical framework with well-established methodologies. An interdisciplinary approach seems to be necessary because the development of such a framework – especially in OPAALS - requires expertise from areas such as linguistics, computer science, social science, cognitive science, and others.

Developing practical knowledge production and management systems in epistemic cultures is a very difficult task. From an implementational point of view, researchers try to build a bridge between formal and informal domains by introducing a semi-formal approach. The majority of applications that have been developed so far are based on the functionalist paradigm, i.e. implemented in a formal domain (Whitley, 1991). The *formal approach* attempts to model a domain using heuristics and rules of thumb and considers the computer system to be the central component that regulates, monitors and guides the flow of information within an organisation. In contrast, *information system* approaches address domain modelling with a rather different focus, viewing the organisational information system as primarily a social system (ibid). The focus of information systems research lies on the system (computer, software or expert systems) as one part of the organisation. Since daily conversations within an organisation are often informally structured, the processes of formalisation and automation are important issues in system design. It is necessary to design software systems that perform these tasks efficiently and user-oriented. However, it is also important to re-emphasise that the computer system is only one component within the formal system which is itself part of the wider informal environment where formal and informal co-exist. This can be achieved through a *socio-technical approach* that views the technical system as only one element in a wider social environment (ibid).

The communication of knowledge within a semi-formal domain is considered to be an important factor in system design. Some parts of the semi-formal domain are the results of subjective interpretation and different users of the system can form different interpretations of this items. As a result, it is no longer possible for the system to control the interaction. The system must be designed as a tool that supports the users in the interaction (Whitley, 1991).

Table 2 summarises the differences between the two approaches, formal and semi-formal, referring to Whitley's “Two approaches to developing expert system” (1991, p. 125).

Table 2. Formal and semi-formal domains (based on Whitley, 1991)

	<i>Formal</i>	<i>Semi-formal</i>
<i>Approach used</i>	Functionalist	Socio-technical
<i>Relationship between system and domain</i>	System models the domain	System is an integral part of the domain
<i>View of reality</i>	Single, measurable	Variety of possible versions
<i>Conceptual view of system</i>	Designed as a machine Machine-centred System controls the interaction	Designed as a tool Human-centred System supports the interaction
<i>Knowledge</i>	Available in discrete chunks	Socially constructed
<i>Involvement of users</i>	Little consideration given	High consideration of users

One of the major problems to move toward a semi-formal approach is the complexity of natural language. The information that is presented in natural language to humans is preferred for its familiarity and expressiveness. In this sense, knowledge about the task and its context will be in a form of semi-formal representations that combine structuring connectors with free text to provide an interface which both allows users to express what they mean and is sufficiently structured to allow machine processing (Chklovski et al., 2005).

One way to approach natural language usage in a formal language environment (i.e. machines) is the introduction of ontologies. The reason ontologies gained a lot of research interest and popularity during the past years is largely due to what they promise: “A shared and common understanding of a domain that can be communicated between people and application systems” (Sure et al., 2003, p. 279). *Ontologies* are “formal specifications of conceptualisations, described by axioms and explicit definitions, which are used for the precise semantic representation of concepts from different areas of knowledge” (Herre & Heller, 2006, p. 108).

Sheth and Ramakrishnan (2003) regard semantics as key ingredient in the Web infrastructure as well as the next generation of information systems applications. The Semantic Web vision is realised through representation, acquiring and utilising knowledge, and semantics is considered to be the best framework to deal with the heterogeneity, massive scale, and dynamic nature of resources on the web. The Semantic Web is one of the web forms “in which information is given well-defined meaning, better enabling computers and people to work in cooperation” (p. 41).

As an ontology represents a part of the domain or the real-world it can be used as a first step in the process of connecting formal and semi-formal domains. According to Sheth and Ramakrishnan (2003), a majority of the Semantic Web applications that have been developed or envisioned rely on three crucial capabilities: ontology creation, semantic annotation and querying/inferencing.

The advantage of applying an ontological approach includes several aspects:

- Community members have a common representation of knowledge;
- a system that provides a common interpretation for the members' interactions, discussions, knowledge and information representations;
- a system forming the basis for defining metadata or semantic annotation.

In the context of the OPAALS community, communication between different domains and the semantic foundation of domain knowledge require a uniform framework based on interdisciplinary ontologies. First steps towards this challenging task have already been taken by, for example, establishing a joint lexicon (JOIE), which is administered by all OPAALS members. It includes terms and concepts with their domain specific explanation and will form the basis for an interactive semantic lexicon. This endeavour combines an interdisciplinary approach with partners from computer science and linguistics working on a cross-disciplinary ontology. Another form of ontology work is being conducted in task 6.7, where a show case example of an ontology is being used in a SME in order to provide valuable test results for an integrative digital ecosystems ontology that spans both research and business fields.

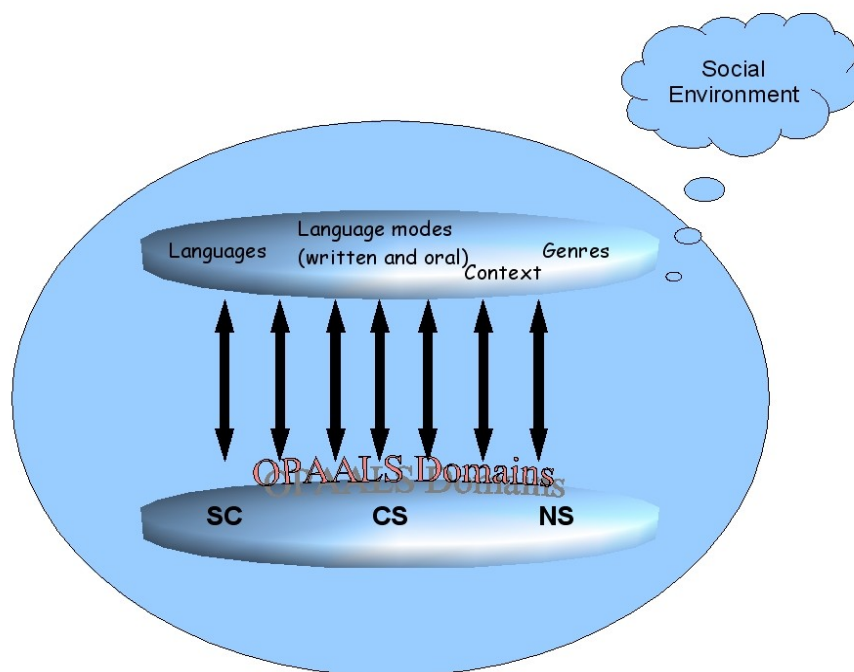
1.3.4) Domain-Specific Languages and Languages for Special Purposes

The difficulty in the development and implementation of semi-formal languages regarding the differences between formal representation and natural language lies in the complexity, variety, and ambiguity of natural language. An effective way to solve this problem is to use approaches from Domain-Specific Languages (DSL) and Languages for Special Purposes (LSP). The general idea is to identify and analyse language in a specific context. 'Context' means either a scientific domain in general (for example, social science, natural science, computer science, etc.), or a particular field/purpose of study (for example, medicine, biology), or even a particular organisational, cultural and social environment.

Referring to the OPAALS community as a subject of research, it is suggested to apply the DSL and LSP approaches in order to develop a solid analytical framework and to apply the results to different research tasks (formal and semi-formal languages, evolutionary framework for language, *metaphorological* tool-kit, etc.) of the project. Although the interdisciplinary, intercultural, and multilingual characteristics of the OPAALS community make the research design more difficult, they also enrich and make it more attractive for other projects and fields of investigation.

Based on research results in the areas of DSL and LSP, several factors can be identified which focus on language and its discourse analysis: domains, languages, language mode (written or oral), genres, context and social environment (Figure 2).

Figure 2. Factors influencing discourse within the OPAALS scientific community (based on Hilbert et al., 2006; Rindflesch & Fiszman, 2003; Auger, 2005)



The OPAALS community consists of a combination of three domains: social science (SC), computer science (CS), and natural science (NS). These domains work, interact, collaborate and communicate with each other in a social environment (the OPAALS community). Most of the processes between and within these domains occur through language(s) in a particular context, and by using different genres or sub-genres (email, journal article, etc.), and of course in different modes (oral and written).

In current studies, two approaches in DSL and LSP can be found that take the *domain factor* into account: Electronic texts from one domain are analysed or texts from two different domains are analysed and compared. For these purposes, different architectures and frameworks have been developed. Hilbert et al. (2006) developed a relational discourse parsing architecture for text documents of a complex text type, namely scientific articles. The documents have been chosen from two domains, psychology and linguistics. The research is based on automatic discourse analysis that deals with less structured texts. Among traditional morpho-syntactic discourse features, the discourse parser is equipped with a *logical document structure* used to build a hierarchy and to

provide the document's graphical layout structure. Additionally, a *genre-specific text type structure* which is based on discourse relations between texts spans on higher levels of discourse. Finally, a *thematic structure* indicates semantic relations between concepts and is represented as a two-level block: the local level where the focus lies on annotating the relations between sequential elementary discourse segments, and the global level which puts complex discourse segments into relation (ibid).

A slightly different approach has been described in the paper “The interaction of domain knowledge and linguistic structure in natural language processing: interpreting hypernymic propositions in biomedical text“ by Rindflesch and Fiszman (2003). The focus of the applied methodology lies on taxonomy and taxonomic relations between different concepts. The authors represent their approach as a way of improving the semantic interpretation in biomedical documents. Similarly to the previous study, this approach is based on two paradigms: the linguistic formalism and domain knowledge. The authors address syntactic structures as being represented by two mechanisms, “a shallow categorical parser and an underspecified dependency grammar” (p. 464). Although the focus of this study lies on the inter-relationship of syntax and domain knowledge in expressing hypernymic propositions, the discourse function of this phenomenon plays an important role. A significant aspect is the distinction between given (or old) information and information that the speaker assumes is being introduced to the listener as new. According to the authors, the hypernymic propositions „provide a means of facilitating the flow of information by accommodating this distinction and can be thought of as definitions embedded in a discourse“ (p.465). The definitional nature of the hypernymic proposition provides a mechanism for serving the same function within a discourse unit, where the specific concept is the new information and the general is the old. In their approach, the authors use semantic types from the Semantic Network approach to ensure that the two concepts involved are compatible. By means of introducing the Metathesaurus – a large repository of concepts, classifications, and coding systems – one is able to determine which concept is more general and which more specific. The architecture is based on the Unified Medical Language System (UMLS) containing three components: the Metathesaurus, the Semantic Network, and the SPECIALIST Lexicon. The SPECIALIST Lexicon and associated lexical access tools provide syntactic information about terms in general and medical English. In Metathesaurus terms that have equivalent meanings are organised into unique concepts, which form the organisational core of this component. Each Metathesaurus concept is also assigned to “one or more semantic types such as Disease or Syndrome or Pharmacologic Substance that categorise concepts in the biomedical domain” (p. 466). The Semantic Network organises these into two single-inheritance hierarchies, one for entities and one for events. In addition, associative relations are assigned between semantic types; these semantic propositions represent knowledge that is accepted as being valid in the biomedical domain.

In relation to the second feature of discourse (which *languages* are used), most of the research studies concentrate on one language only. Rindflesch and Fiszman (2003) provide an architecture

that encodes hypernymic propositions through the interpretation of the linguistics structures and is based on biomedical documents written in English. Another example is a SACOT – Semi-Automatic Construction of Ontologies from Texts – which deals with domain ontologies and concepts based on English electronic texts (Auger, 2005). In contrast, the study “A text-technological approach to automatic discourse analysis of complex texts” (Hilbert et al., 2006) uses a large corpus containing 183 scientific articles written in English and German as the input of the parser described previously.

In relation to *genres*, there is a variety of studies using scientific articles as an object of research. Some of them compare two different genres or sub-genres. Hilbert et al. (2006) use electronic texts as input for the parser and differentiate between two sub-genres: experimental and review. In the paper “Applying natural language processing techniques to domain knowledge capture“, Auger (2005) analyses electronic texts and their semantic capabilities. The study deals with the question „how natural language processing techniques can support and automate ontology engineering process and enable capture of domain knowledge objects from electronic texts“ (p.1). In order to analyse the semantic capabilities and “to make sense out of all available information” the author proposes the use of the domain ontologies approach and introduces a framework to investigate, develop, and validate natural language processing. SACOT „approaches as scientific means to capture knowledge objects contained in open source electronic texts and turn them rapidly into broad domain ontologies to be used in third-party applications“ (p.1). This framework is based on domain ontologies and as a result on concepts and relations among these concepts. The author points out that the major disadvantage in the existing information retrieval systems is its focus on statistical methods and therefore their limited precision in the capturing of domain-specific knowledge objects. Using “low level linguistics” methods such as stemming or part-of-speech tagging does not provide an access to any form of “prior semantic knowledge”. To avoid this limitation, three Natural Language Processing (NLP) techniques have been investigated within the SACOT framework (Auger, 2005):

1. Terminology extraction techniques
2. Named entities extraction techniques
3. Semantic relations extraction techniques

Auger argues that the SACOT framework’s originality and advantageousness lies in the fact that “all its internal processes are based on the exploitation of prior linguistic knowledge (domain-specific texts, dictionaries, taxonomies, contrastive datasets, semantic relation markers, explicit named entities) taking extensively advantage of recent NLP techniques” (Auger, 2005, p. 2).

From the “*language modes*” perspective most of the current studies in the fields of DSL and LSP deal with electronic forms of written language. Rindflesch and Fiszman (2003) use electronic

documents from the biomedical domain in order to identify and interpret hypernymic propositions “in which a more specific concept is in taxonomic relationship with a more general concept” (p. 462). Hilbert et al. (2006) use electronic texts as input for the parser.

This reductionist focus on electronic texts may negatively influence the quality of an overall analysis of the discourse organisation in epistemic cultures. Therefore, an additional feature needs to be taken into account: *language mode* (written and oral language). Based on the elaborations above regarding context and socio-cultural influences, it can be argued that the *context* of the particular action or situation and overall social environment affects the choice of lexico-semantic features within the OPAALS community.

In the context of language and discourse, one of the most important questions is „How can people in fact maintain a consensus about what they mean?“ (de Beaugrande, 1997, p. 11). The explanation developed by de Beaugrande is that people use and understand language precisely because their language system and knowledge about the world are constantly interacting with each other. This system consists not only of the rules applied by the speaker, but also strategies of selecting rules and adapting them to the context. “Alone, the rules would either be too general to provide criteria for quick selection and application, or else too specific to be usable on a wide range of occasions“ (p. 11). The language is usually used in discourse with others and is never the whole language of the interaction participants. It is not even the whole knowledge of the language, it is rather the currently active version of it (ibid). Our choice of words and their usage may be conditioned by what we are talking about and to whom we are talking (Farghaly & Hedin, 2003). Signs and their meaning are formed by social groups (i.e. discourse communities) primarily as part of the social division of labour in society. Such discourse communities may be very different in size and structure. A large number of groups may develop symbol systems and share knowledge which they do not share with the rest of society. There may be a considerable degree of common knowledge and shared meaning (Hjorland, 2002). Additionally, the social level of the language domain analysis would cover the variability of lexical item choice in response to the different social parameters (social status, education, age, sex, etc.). According to Baumann (2001), the selection of lexico-semantic items depends on the social relations among communicative partners.

In the paper „Discourse communities and communities of practice: On the social context of text and knowledge production“ Pogner (2005) discusses two concepts of 'community' (discourse community versus community of practice) from the viewpoint of text production, social learning, information management and knowledge sharing. The study consists of two building blocks. Firstly, it introduces the theoretical exercise that tries to determine „the degree to which the concepts of discourse community and community of practice are suitable for investigating the social and organisational context of text and knowledge production“ (p. 1). Secondly, the study analyses the contribution of the two concepts to the investigation of knowledge and text production (and of their possible interplay) in organisations.

According to Pogner, communities are characterised by the development of certain behavioural rules and norms, as well as by the feeling of belonging and engagement: “That feeling is a result of shared ways of behaving and thinking – whether it is due to similar interests, tasks, professions or lots. [...] In both cases knowledge is developed and shared” (Pogner, 2005, p. 4). In his research, Pogner differentiates between hard and soft knowledge: hard knowledge is defined as declarative knowledge that “can be articulated and codified, is learned via the procession, evaluation and revision of various documents” (p. 11). Soft knowledge is characterised as a “non-codifiable, procedural knowledge, which is not so easy or impossible to quantify, grasp and store, can only be acquired in processes of participation (membership, activity, interaction etc.) and reification (shared documents, texts, instruments and other artefacts)” (p. 11).

Pogner identifies several characteristics of discourse communities and communities of practice based on the concepts of knowledge and language/text. A summarised overview of them is represented in Table 3.

Table 3. Discourse Community and Community of Practice

<i>Discourse Community</i>	<i>Community of Practice</i>
Characterised by the mutual dependence of language use, membership community and various specific patterns of language use.	Characterised by sharing stories, cooperation and a form of social construction of the community of practice.
The group shares more than a particular native tongue or symbol manipulating skill.	Includes different workplaces and companies but especially those organisations that want to / have to act in the context of the knowledge economy.
People feel connected by virtue of their shared discourses and the work the discourse enables them to do.	Combined learning in the sense of an activity (practice) with learning in the sense of belonging (community).
Members of the discourse community use their texts in order to demonstrate their membership in the community to which they belong or wish to belong.	
Shared language and discourse as well as shared practices of thinking, research and learning.	
<i>Are characterised by loose connections between the members and by higher degree of informal self-management.</i>	

Many of these discursive practices which occur in both types, discourse communities and communities of practice, take place in electronic newspapers, on electronic conferences, in electronic discussion lists, web logs etc. using ICT. Pogner pointed out that in both types domain-specific texts, discourses, and domain-specific language play an important role. However, “with communities in which the discursive element is predominant, written texts occur more often than they do in communities in which the practical element and the spoken language are predominant” (Pogner, 2005, p. 10).

Predominantly, case studies related to this research refer to internal and external communication at various workplaces: “Text production by consulting engineers in the preparation of an energy concept (project team), semi-public email communication at a university department (department) and the work of two project teams developing IT solutions for a bank (project team)” (Pogner, 2005, p. 3). Since the use of ICT plays an important role in the case studies, the author takes the concepts of virtual teams and distributed communities into account too. He notes that the data consist of written documents (mainly in electronic form), qualitative interviews and ethnographic data such as semi-detached or participating observations. Pogner (2005) noticed that not only the text production, commentary and revision, but also visits, meetings, phone calls etc., contribute to knowledge gathering and solution strategies. Additionally, Pogner found out that the text helps to focus meetings and discussions, contributes to organising, planning and coordinating the task at hand, and it functions as a communication tool. He argues that the challenges which the participants face in interaction with other experts and clients via different media, genres, and forms of communication (face-to-face, telephone, fax, email, discussion protocols, text drafts and versions, drawings, viewings, inquiries etc.) cannot only be explained by using the concept of the community of practice; some aspects can be dealt with more adequately by referring to the concept of discourse community. The communities have a strong effect on how texts and knowledge are produced and passed on in organisations or departments and project teams within them – or at least how this could happen. Collaborative text production and/or other collaborative practices play a constituent role for the existence of communities which can promote knowledge production and learning in organisations. Summing up, discourse communities and communities of practice are important components of the social context of the production of texts and knowledge, even if they are not the only existing components.

Marshak and Heracleous (2005) explore the notions of discourse by introducing concepts of *organisational discourse* and *action research*. They depict a conceptualisation of organisational discourse as situated symbolic action. This approach is supported by introducing the analysis of a meeting of senior managers during an organisation development intervention. This perspective “encourages a more holistic understanding of organisational contexts and offers an actionable framework to help make sense of workplace episodes and choose appropriate interventions” (p. 69). The specific context for the discussion and practical interpolation is a meeting of the top management of a high-tech organisation to decide on a new business model. According to the

methodological description of this study, one of the authors participated in the meeting in the role of the organisation development consultant.

In the context of the theoretical discussion and background, the authors present the notion of organisational discourse and identify it as „an eclectic variety of approaches based on a range of disciplines where the central focus is the role of language and linguistically mediated experience in organisational settings“ (Marshak & Heracleous, 2005, p. 70). The innovative idea of using different aspects of discourse, text or other media in studies in context of managerial and organisational phenomena provides a new dimension and possibilities of research.

1.4) Summary

So far, different aspects of discourse organisation including theoretical backgrounds and approaches of research were introduced. Although language is the focus of analysis, the importance and necessity of exploring the notion of culture needs to be recognised and to show how the concepts of knowledge, language, communication, and social environment are “living together” and can be applied to the different research areas and domains existing within OPAALS.

The aforementioned empirical study (see also chapter 4) will lead to publications focusing on digital ecosystems and epistemic cultures, digital ecosystems and the role of language, and on-line communities as networked organisations. The language and culture focus also connects to task in work package 7, discussing notions of community from a formal point of view. Connections and further collaborations have already been discussed with work package 5, where accountability, security, and trust are addressed from a computer science point of view and offer synergies with a social science approach to community building (e.g. regarding different concepts of trust and security within an on-line community). The strong linkage to work package 10 has been mentioned before and is also executed by its connection to deliverable 10.5, which includes a sound discussion on 'knowledge'. Finally, chapter 1 represents the broadest and most detailed part of this deliverable as it also serves as a basis for task 6.6 “Evolutionary framework for language”, considering the notions of context, ontology building and more generally dynamic genres¹ and discourse models. Specifically the computational linguistic parts in this deliverable point to an integrative framework for natural and formal languages, which represents one of the major challenges in OPAALS. Moreover, work conducted in work package 2 discussing a formalised language processing framework (SBVR) from a linguistic point of view is to be further elaborated in Phase II within work packages 2 and 6.

¹ 'Genres' represents an important methodological and theoretical aspect of communities and knowledge production. Therefore, this field is elaborated in detail in the following chapter.

2) Epistemic Cultures

2.1) OPAALS and the notion of epistemic cultures

There is a variety of characteristics of epistemic cultures that can be applied to OPAALS. According to Evers, the “culture of organisations” is turned into an *epistemic culture*, a culture of knowledge production and utilisation. Individuals are no longer viable as epistemic subjects, but have become integrated into the gigantic “laboratory” of the “learning organisation”, creating and absorbing knowledge (Evers, 2005). Technical, social and symbolic dimensions of intricate expert systems are combined into the epistemic machineries of science research. The boundaries of epistemic cultures are right across the sciences in general (Evers, 2000). Within OPAALS these boundaries exist across three scientific domains: social science, computer science and natural science.

The theoretical framework and definition of epistemic cultures was developed, among others, by Knorr-Cetina. She defines epistemic cultures as “those amalgams of arrangements and mechanisms [...] which, in a given field, make up how we know what we know. Epistemic cultures are cultures that create and warrant knowledge [...]” (Knorr-Cetina, 1999, p. 1). Likewise, according to Evers (2005), the boundaries between knowledge and society are blurred and epistemic cultures are complex blobs of knowledge, actions and emotions.

Evers (2000) defines the following characteristics of epistemic cultures:

- Independence from the outside control;
- collaboration and at the same time competition of resources, recognition, and excellence;
- existence of behavioural rules, “but no undue regulation of values or prices”;
- high level of competition (without open conflict);
- a high degree of autonomy of decision making.

The term 'epistemic' is related to the concepts of knowledge creation, production and sharing. One important goal of OPAALS is to find out how knowledge as a relatively abstract concept and “OPAALS culture” involving real people, relations and communicative processes between them are correlated and “live together” within a/the digital ecosystem. The geographically distributed character of the OPAALS research community that links people across countries and organisational units provide an additional dimensional complexity of the epistemic culture concept. Factors such as distance, organisational affiliation, cultural differences, language, etc. challenge the process of community building and evolution. Cultural differences can easily lead to communication difficulties and to misinterpretation; language may also introduce a very basic barrier of communication. Since the major production subject of the OPAALS community is *knowledge*

(related to different aspects, fields, and dimensions of research), it is important to understand the role and place of it within culture and community.

McCarthy (1996) suggests two recurring themes that summarise the intellectual tradition of the sociology of knowledge. They represent two seemingly divergent ideas about knowledge's place within a social order. The first proposition is that *knowledge is socially determined*. This idea has dominated the sociology of knowledge since its inception. The second proposition, *knowledge constitutes a social order*, asserts that knowledges are not merely the outcome of a social order but are themselves key forces in the creation and communication of a social order (ibid). The role of knowledge within a community can be described as a major factor of production parsing capital and labour (Evers, 2000). Universities, research centres and institutions are important organisations of knowledge production. Based on this concept, Evers pointed out several differences between knowledge and other factors of production that are also relevant in the context of the OPAALS project:

- Knowledge is difficult to analyse, investigate and measure;
- once knowledge has been produced it can easily be reproduced or copied;
- transaction costs in trading knowledge are low;
- knowledge experiences rising marginal utility. The more an expert, a group of consultants or an organisation know, the more valuable become individual pieces of knowledge (Evers, 2005, p. 7).

According to Knorr-Cetina (1999): “A knowledge society is not simply a society of more experts, more technological gadgets [...] It is a society permeated with knowledge cultures [...]” (p. 7). Hence, a knowledge society can have the following characteristics:

- Higher educational degrees of the society members (=knowledge workers);
- its industry produces products with integrated artificial intelligence;
- development towards the intelligent organisations;
- knowledge is organised in the form of digitalised expertise, stored in data banks, expert systems, organisational plans and other media;
- existence of centres of expertise and a polycentric production of knowledge;
- epistemic culture of knowledge production and knowledge utilisation (Evers, 2000, p. 8).

The emergence of a productive epistemic culture (of knowledge production) is difficult to achieve. According to Evers (2000), the preconditions for the development and growth of epistemic cultures and their shape and contents should be investigated and understood to explain the morphology of knowledge production. Knorr-Cetina (1997) states that a knowledge society is not simply a society of experts or an increasing production and flow of knowledge, but rather a society into which knowledge cultures have spilled and woven their tissues into society as a whole.

Knorr-Cetina (1997) emphasises that the most profound changes in modern societies is the expansion of object centred sociality. In contrast with a variety of other social science studies that exclusively focused on human bounds formed through normative consensus, she argues that integration based on knowledge objects may create communities 'in thought' (ibid). Like many sociologists of scientific knowledge she takes her examples from the (natural science) laboratory focusing on how scientific research works in practice. She points out that such knowledge settings are no longer limited to science. Over and above, Knorr-Cetina describes knowledge objects as the goal of expert work and also what they are interested in. Experts' relationship to knowledge objects is however, not only emotional bounds, it is also dynamic and ambivalent (Knorr-Cetina & Bruegger, 2002).

In reference to the Network of Excellence and concept of epistemic culture, we identify additional factors forming the OPAALS community and influencing its growth and development: Interdisciplinarity, open knowledge community, autopoiesis, and reflection. These factors are twofold:

- On the one hand, they are the settings need to be taken into account in the research fields existing within and across three scientific domains: Social science, computer science and natural science.
- On the other hand, these factors reflect the uniqueness and complexity of the OPAALS community itself.

2.1.1) Interdisciplinarity

The OPAALS community is a system of three interacting and collaborating domains: Social science, computer science and natural science. One of OPAALS's overarching aims is to build an interdisciplinary research community in the field of digital ecosystems and therefore focuses on knowledge production and management. Besides the dispersed organisation of the project participants and institutions, the members' affiliation with different research fields within and/or across scientific domains provide a challenging background for the analysis of social and communication network specifications, knowledge production and sharing as well as aspects of language and knowledge representation.

2.1.2) Open Knowledge Community

Another goal of the OPAALS project is to build an Open Knowledge Community. Important factors the project deals with are *relevance* and *accessibility* of the content. The purpose of the OPAALS community is to promote both the creation and the exchange of this content within a digital ecosystem as well as to make the existing and newly created knowledge accessible, reproducible and re-usable. The content could have many different forms like text, picture, diagram, map, etc.

Talking about OPAALS as an open knowledge community, we introduce two perspectives of “openness”: social and technical. *Social openness* consists of ensuring that a work is made available by supporting sharing and reuse as well as collaborative working processes within and outside the community. *Technological openness* means that there must be an agreed form and format of the knowledge and information as well as a possibility to access not only at the level of machines (computer, databases, etc.) but also at the level of human interaction and collaboration. These forms of openness could be the first steps towards the Open Knowledge System (OKS) development within the OPAALS project. Through socio-technological processes the community members are able to disseminate their thoughts, ideas, information and knowledge not only to their colleagues and collaborators, but also to the rest of the world. There is a possibility of involving the “outsiders” to participate in processes of collaboration within and outside the OPAALS community. In this context, the processes of knowledge dissemination and collaboration are available through the open access to the data.

2.1.3) Autopoiesis

The concept of autopoiesis comes from biology and was originally attempted to characterise living systems as being auto-productive. The notion of autopoiesis is often associated with the term “self-organisation”. The main question is how to apply the concept of autopoiesis to the human societies in general and to the OPAALS project in particular. As society is constituted of living organisms (humans), the biological phenomenology could also be applied to our society. Referring to the notion of epistemic cultures, we might see the knowledge in the organisational or cultural context as a survival value of the community and language as the catalyst of its development. Allott (2001) identifies language as a „missing link between genetic and cultural evolution”. The great feature of language is its ability to create different structures through modelling unobservable features of the universe and to generate the sciences. We are able to combine, share, and manipulate these features (Allott, 2001). The social significance of language, writing, printing, data storage etc. are significant factors in symbolical knowledge representation and providing of objectified knowledge. In this sense, knowledge can be seen as an universal phenomenon, or an anthropological constant.

2.1.4) Reflectivity

Reflection is an active process through which we curiously witness our own experience and take a closer look at it and explore it in depth. The key issue of the reflective process is the observation of one's own actions, experiences, ideas, knowledge and to examine them through a critical lens. The ability to explore and be sceptical to some extent opens to the community members the new dimension of research and knowledge production. In this context, we also introduce the notion of reflectivity as an application of the OPAALS research questions to the community itself. From this perspective, we have the ability to see people in two different situations, as developers (designers and/or researchers) and as users (and / or object of research). This allow us to gain the results and to learn through experience.

2.2) Knowledge cultures

There are different aspects of research in the field of knowledge cultures: Epistemic cultures, communities of practice, epistemic communities, organisational communication, genre systems and media richness. Independently from the term used – epistemic culture, community of practice, epistemic community, or just organisation – all these concepts have two important issues in common: *language* as a catalyst of development and *knowledge* as a product. Language influences the structure of an epistemic culture, its development, and dynamics. In society language plays a dominant role as people exchange ideas, interact, and create new knowledge through language. Language supports and makes possible both the development of the individual and social consciousness (Allot, 2001).

2.2.1) Epistemic cultures

Knorr-Cetina proposed in her book “Epistemic cultures” (1999) the notion that different scientific fields exhibit different epistemic cultures. In her study she focused on the comparison of two different disciplines: High energy physics and molecular biology. In high energy physics “individual persons are displaced as epistemic subjects in favour of experiments themselves” (Giere, 2002, p. 637). In the molecular biology field, Knorr-Cetina identifies a two-level structure of a laboratory:

1. The lower level that consists of researchers each working on their own project.
2. The upper level which is the whole laboratory usually managed by a single person (director).

The individual nature of the lower level seems to be important and interesting when comparing two epistemic cultures. The difference between the two cultures lies in the observation that in the biological laboratory, the person is an epistemic subject. The laboratory itself, procedures, experimentations, and objects achieve their identity through individuals (Knorr-Cetina, 1999).

Inspired by the theory of knowledge society, Evers (2005) applies the notion of epistemic culture to the market and economic development. He argues that knowledge gained importance in economic development in the last years, replaced industrial organisation and production, and became the major source of productivity. “In fact, the largest share of value added in modern computer technology does not rest on the value of the material used or the input of labour and capital, but on the knowledge embedded in the final product” (Evers, 2005, p. 6). In modern society, knowledge becomes one of the most important factors of production, overtaking capital and labour. According to Evers, “to achieve the status of a knowledge society, it is not enough to buy and to consume knowledge, but also to produce it” (p. 10). In contrast to Knorr-Cetina, whose focus lies on the construction of the knowledge machineries, Evers concentrates on knowledge creation. In the context of epistemic machineries, he identifies three dimensions: technical, social and symbolic. The author expands the idea of epistemic cultures from Knorr-Cetina and argues that epistemic cultures exist not only in the “laboratories of natural science research, but are institutionalised in various ways in the New Economy of globalised knowledge societies” (Evers, 2005, p. 11). The institutional aspects or contours of epistemic cultures are based on the size (“there has to be a sizeable number of persons”), control dependency (“who are relatively independent of outside control”), distribution or community network (“who work closely together”) and competition among members (“who are pitted against each other in competition for resources, recognition and excellence”) (p. 12).

Additionally, Evers (2005, p. 12) proposes the idea of resemblance between an epistemic culture and culture of markets:

- “there are stringent rules of conduct”;
- “no unique regulation of values or prices”;
- “there is competition but no open conflict”;
- “there is a high degree of autonomy and decision making”.

Smeby (2006) narrows the notion of epistemic culture and its dynamics of knowledge production to professional development. He examines the “relationships between initial education and epistemic strategies among beginning professionals (nurses, teachers and social workers)” (Smeby, 2006, p. 3). The author argues that a greater emphasis on knowledge and epistemic cultures is needed to understand the challenges of professionalism in modern knowledge societies. In the context of knowledge production and epistemic cultures, professionals do not “just need knowledge as a basis

for carrying out their tasks; they also have to a much greater extent to defend their professional practice scientifically towards other professional groups as well as the lay audience” (Smeby, 2006, p. 4).

Smeby (2006) points out several important issues that can be linked to the research of the epistemic cultures within the OPAALS framework:

- The differences between epistemic cultures;
- the extent to which individual characteristics are related to epistemic strategies;
- the relevance of abstract codified knowledge;
- the analysis of acquisition from a sociological perspective (“...focusing on how professionals link to more abstract modes of knowledge and get connected to knowledge”) (Smeby, 2006, p.6).

The research results indicate that both perspectives focusing on acquisition as well as perspectives focusing on participation are important to shed light on professional development and learning among beginning professionals.

2.2.2) Epistemic communities and Community of Practice

Epistemic communities can be defined as small groups of agents working on a commonly acknowledged subset of knowledge issues and “who at the very least accept a commonly understood procedural authority as essential to the success of their knowledge activities” (Cowan et al., 2000, p. 227). Epistemic communities can thus be defined as a group of agents sharing a common goal of knowledge creation and a common framework allowing to understand this trend. Within an epistemic community, agents are bound together by their commitment to enhance a particular set of knowledge. It can be defined as “a network of professionals with recognised expertise and competence in a particular domain and an authoritative claim to policy-relevant knowledge within that domain or issue-area“ (Haas, 1992, p. 3). Members of an epistemic community may have different academical, educational or professional backgrounds. However, according to Haas (1992), they must have several aspects in common:

- 1) “a shared set of normative and principled beliefs, which provide a value-based rationale for the social action of community members”;
- 2) “shared causal beliefs, which are derived from their analysis of practices leading or contributing to a central set of problems in their domain and which then serve as the basis for elucidating the multiple linkages between possible policy action and desired outcomes”;

- 3) “shared notions of validity-that is, inter-subjective, internally defined criteria for weighing and validating knowledge in the domain of their expertise”;
- 4) “a common policy enterprise- that is, a set of common practices associated with a set of problems to which their professional competence is directed, presumably out of the conviction that human welfare will be enhanced as a consequence“ (Haas, 1992, p. 3).

Another concept existing in the context of knowledge and culture is the notion of 'communities of practice' (CoP). In CoP individuals accumulate knowledge in function of their own experience. This might correspond to the reflective character of the OPAALS research community. According to Baumard (1999), the quality of this knowledge within CoP depends on two factors:

- The variety of individual experiences in interaction;
- knowledge of the experience.

The first task of such communities is thus to create a ‘codebook’ as knowledge circulating within epistemic communities is explicit but not codified since it remains mainly internal to the community (Baumard, 1999).

2.2.3) Organisational Communication

Communication has nowadays become one of the most important assets in organisations. This is because organisations cannot be considered just as “containers” of individuals with common aims, but should to be regarded as evolving social contexts in which real persons face various situations and problems. Thus, communication is the means they possess in order to understand and adapt to the dynamics of these changing environments (Cuel & Ferrario, 2006).

Before introducing studies and papers related to organisational communication in the context of epistemic cultures and their discourse organisation, two important concepts – organisation and communication – should be defined. Miller (1995) defines organisation in terms of five features:

1. Social collectivity: An organisation consists of two or more people. Even at the simplest level of organisation, social collectivity requires “some level of understanding, coordination, and communication to act together” (Miller, 1995, p. 6).
2. Goals: The definition of organisation from this point of view assumes the multiplicity of (shared) goals the members of an organisation have.
3. Coordinating activity: This feature addresses the process through which goals are achieved by the social collectives.

4. Structure: This feature is seen as patterns formed through coordinating activities.
5. Environmental embeddedness: This feature is based on the “belief that no organisation exists in a vacuum” (Miller, 1995, p. 9). Thus, all organisations exist within a social environment that influences the organisational activities.

The research field of organisational communication has the twofold purpose of “understanding how communicational processes shape organisations”, and “understanding how organisational life influences the form and content of communicative acts of the individuals who interact within it” (Cuel & Ferrario, 2006, p. 6). Based on studies in the field of organisational communication, four dimensions of research can be identified:

1. The dependency between communication processes and different types of organisation (Putnam et al. 1996, Cuel & Ferrario, 2006).
2. The question of influence: organisational hierarchies influence communication processes, style, flow, etc. Correspondingly, communication influences the structure and work flows in organisations (Putnam et al. 1996, Cuel & Ferrario, 2006).
3. The role and usage of ICT (e.g. CMC) in epistemic or organisational environments and regarding the choice of different communication channels (Scott & Timmerman, 2005).
4. The importance of coordination processes in the work flow at different hierarchical levels of organisations (Cuel & Ferrario, 2006).

Davenport et al. (1998) point out that the processes of communication in an organisation should be shaped according to the information strategy existing within the organisation and identifies four different models:

Information anarchy, where coordination and communication processes depend on the ability and professionalism of a single worker who manages her personal information and networks of communication channels.

Information hierarchy: The twofold communication mode (top down and bottom up) between members of an organisation is adapted to communication channels that are organised by the management.

Information feudalism: “Workers share knowledge within each single unit, teamwork, or community. Communication channels are organised and managed within each single group and no official communication channels occur across units” (Cuel & Ferrario, 2006, p. 3).

Information federalism: The personal, informal, and formal communication forms that occur inside a team or organisation are placed side by side.

These four models of information strategy are often represented in organisations together in order to satisfy all organisational needs.

Regarding the technological impact on organisational communication, Scott and Timmerman (2005) explore three issues regarding the use of multiple workplace communication technologies:

1. The relationships between distinct forms of apprehension and acceptance/usage,
2. the relative contribution of computer-mediated communication (CMC) apprehension for predicting usage of communication technologies, and
3. changes in these relationships over time.

Related to the already mentioned areas of research, several questions are of interest. For example, it is important to explore the anxiety towards technology in general, and computers specifically: “Because many of the new workplace technologies are primarily for communication purposes (e.g., e-mail, video conferencing, voice mail), users’ aversions to communication should influence the likelihood of using new communication technology” (Scott & Timmerman, 2005, p. 684). Therefore, it is important to find out how both types of communication, oral and written, and computer apprehension “predict usage frequency for a variety of new technologies” (p. 685).

Additionally, Scott and Timmerman (2005) argue that several distinct apprehension types could help explain communication technology use at work. For this purpose, they suggest using the combination of apprehension measures, which “may prove useful as an even more specific way to examine this issue” (p. 685).

Finally, the usage of CMC is an important factor in the organisational communication theory framework. According to the authors, by using CMC in the work place people “may experience a sense of anxiety about using computers to communicate” (Scott & Timmerman, 2005, p. 685). To examine this possibility, they measure CMC apprehension and evaluate “the predictive power of this variable relative to distinct apprehension types in explaining workplace technology use” (p. 685).

A different approach has been introduced by Cuel and Ferrario (2006). Their research is based on a case study in which the authors have “adopted the ethnomethodological approach to analyse the impact on the internal communication processes of the implementation of a KM system, called Sispes, in a residence for elderly (Italy)” (Cuel & Ferrario, 2006, p. 1). Sispes is a web based platform which supports communication processes and knowledge management according to a customised work flow management system. The authors investigate the evolution of the communication processes within the organisation, and the impact of Sispes on knowledge management assets. In the context of their research, the authors distinguish between three types of communication processes: normative, descriptive and constructive. The analysis and adaptation of

the Sispes system serve the goal of understanding how the introduction of it has affected the information strategy of the firm, the workers' attitude, the communication and knowledge sharing processes. Additionally, they emphasise several characteristics of communication:

- Communication processes are unidirectional, directed from the 'top' (the management) of the organisation down to the workers. In this case the purpose of communication is providing a set of rules and requirements to be met. Communication is seen as the carrier of the judgement on what is right or wrong, good or bad.
- Communication processes are bidirectional; the understanding of the real processes and communication within an organisation is more important and effective than the simple act of imposing rules. It allows the management to gather useful information in order to improve the work quality. In this case, communication does not have the purpose of transmitting values, but rather of drawing a faithful description of the organisational reality 'as it is'.
- Communication processes are informal, “horizontally and transversally oriented” processes, meaning that the participants, through negotiation and communication, are able to build the socially constructed concepts and use them inside the organisation.

The basic idea behind this approach is that the organisational reality is not detached from the individuals who perceive it, nor neutrally given (and transmissible); the knowledge of the reality is filtered “by the interpretative schemas of subjects” (Cuel & Ferrario, 2006, p. 10). Communication is seen as a process of creating “new ways of experiencing the reality and from these new social entities can emerge”, rather than a process of exchanging information (p. 10).

The technology introduced in this research manages most of the information flows among social assistance, sanitary, entertainment, and administrative units. In addition to the authentication process and groupware application it provides a complex system of work flow and document management that allows users to add and retrieve information from the corporate archives. According to the authors, the main characteristic of Sispes is that it is focused on the residence's guests and their needs, therefore all information is managed according to this perspective. Regarding *informal communication*, Cuel and Ferrario (2006) identify two subcategories: Completely unstructured, like occasional conversations and gossips, and semi-structured. Unstructured informal communication is very common in work settings. In the context of the Sispes applications, informal communication serves “the purpose of filling the gap determined by the absence of formal communication and information procedures involving all the people working with the same guest but at different times or under different perspectives” (p. 15). The semi-structured communication was enabled through implementation of “Guest diary”.

The second main category of communication, *formal communication*, occurs during (through) different kinds of meetings: Consignation meetings, meetings for the Individual Assistance Plan,

meetings of the units, etc. The proposed approach – classification of different communication processes – provides “a line of interpretation that allows the joint analysis of communication processes and the organisational reality in which they take place” (Clue & Ferrario, 2006, p. 20). The major focus of this study lies in the relation between communication processes and ICT applications.

At a more abstract level, the analysis of this case study has shown that technology is in fact a major social actor, as it modifies pre-existing practices, like in the case of meetings (where the reports of Sispes often give the directions of the discussion) and offers new possibilities, like that of inferring new information from statistical results extracted from the data stored in Sispes.

Marshak and Heracleous (2005) examine the organisational development through the lens of discourse, namely different types of discourse: Discourse as action, discourse as situated action and discourse as symbolic action: “The analysis is also informed by knowledge of the interactional and organisational contexts acquired through the intervention experience” (p. 79).

1. Discourse as action: The analysis of discourse as action focuses primarily on “who said what and what they seemed to be overtly intending to achieve” (Marshak & Heracleous, 2005, p. 80).
2. Discourse as situated action: In viewing discourse as situated action the researchers added several frames of context to “more fully understand the actors’ intentions and their effects”. Besides the interactional context of the meeting, they introduced “the organisational context and the broader industry context” (Marshak & Heracleous, 2005, p. 80).
3. Discourse as situated symbolic action: At this level of analysis, the issue of power takes place in the course of communication processes and collaboration among meeting participants. The authors observed an interesting label called “participant-led” that was a focus of debate, “not only because it would influence existing power arrangements by symbolising who would have the power to control the sales and delivery process, but because it summed up and evoked in a single word the entire contest between old-timers and newcomers over the future business model and culture [...], including all the associated thoughts and feelings of the involved participants” (Marshak & Heracleous, 2005, p. 81).

This study shows that conceptualising and analysing discourse as situated symbolic action has a number of significant, interrelated implications (Marshak & L.Heracleous, 2005):

- It can help to respond to some of the key challenges facing the field of organisational discourse.
- It encourages a more holistic and discourse-sensitive understanding of empirical contexts for organisational researchers.

- Viewing discourse as situated symbolic action offers an actionable framework to organisational actors and organisation development practitioners for making sense of workplace episodes and selecting appropriate interventions.

2.2.4) Genre Concept and Media richness

Common concepts related to epistemic cultures are “epistemic community”, “community of practice”, “knowledge cultures”, etc. The distinctive characteristic of these approaches is twofold: media use is a case in point; academic discourse (genres) another.

The first approach deals with the question why people use different media to exchange different information. As theoretical framework it serves the conceptions of *media richness* and *social presence*. They postulate that a good match between the characteristics of the medium and the topic of communication leads to a good communication. A good fit between medium and task means that a lean medium is chosen for an unequivocal message and a rich medium for a more complex one. If an unsuitable medium is chosen in a communication process, a mismatch occurs. Both models are based on “objective” characteristics of media. According to Rice (1993), the social construction of such characteristics seems to play no significant role (see also Bouwan & van de Wijngaert, 2002).

The social presence and media richness provide important models for prediction of media choice and usage behaviours. In this context, sociability is an important factor in communication and different communication media support different levels of social presence (Campbell, 2006). There is a general consensus within the literature that as bandwidth narrows, the communication channel becomes less suitable for complex social interactions and interpersonal communication (Rice & Williams, 1984). Daft and Lengel (1986) differentiate between rich and less rich media: Media capable of sending "rich" information (for example, face-to-face meetings) are supposed to be better suited to equivocal tasks where multiple interpretations of available information exist. Less "rich" media (= computer-mediated communication) are supposed to be best suited to tasks of uncertainty with a lack of information.

Point of departure of the *media choice* model is that organisational success depends on the ability to choose the appropriate media to avoid equivocal messages. The media choice model postulates that the “subjective” characteristics of a medium are socially constructed. Social influence in the model is thought to consist of direct statements by co-workers, vicarious learning, group behavioural norms and social definitions of rationality (Bouwan & van de Wijngaert, 2002).

The second approach deals with different *genres* and sub-genres of academic writing and how they reflect the way scholars produce, evaluate, and present evidence and their knowledge. Discursive and publication practices are shaped by the way the academic reward system is operationalised from one discipline to another (Cronin, 2003). A great impact on the notion of genre and genre systems

has been made by Yates and Orlikowski. They explore these terms from the organisational standpoint. The collaboration and communication processes are organised and defined by the genre systems. Therefore, examination of the genre system in real organisational environment could lead the research not only of organisational communication but also of knowledge societies (epistemic cultures) to the identifying the features of collaboration processes in general and distributed, computer-supported collaboration in particular.

Talking about genre as a concept in the context of knowledge and knowledge culture, we identify four approaches (directions of research): genre, genre systems, genre taxonomy, and genre analysis.

Genre as a generic term

Formally, a *genre* can be defined as a taxonomic collection of speech or written text types showing the systematic use of genre elements (Lacey, 2000). The genre elements are related to either form or content. Some researchers define the communicative purpose as a fundamental property of a genre (Yates & Orlikowski, 2002). Here, a genre is a template for action enacted within a community to accomplish a socially recognised purpose. This definition is based on two implications:

- Community ownership, which means the conventions of form, contents and purpose reflect the recurrent practices of a community and not individual or opportunistic motivations (Bergquist & Ljungberg, 1998).
- The notion of organisational communication genre, i.e. genre may be applied to the organisational context. This type of genre can be defined as a typified communicative action invoked in response to a recurrent situation within an organisation (Yates & Orlikowski, 1992).

Yates et al. (1997) used the concept of genre in order to examine the use of Team Room, a collaborative application designed to support teams within organisational settings, to facilitate collaborative work in one organisation. The methodology of this study is based on the communication analysis within three teams existing in Team Room over seven months. All teams use genres and genre systems such as meeting documentation, collaborative repository, and collaborative authoring in different ways dependent on the team size, task, and orientation towards the new technology. “The three genre systems observed in Team Room demonstrate different forms of electronic collaboration that both build on and vary from collaboration in traditional media” (Yates et al., 1997, p. 1). In their exploratory field study, the authors used the genre approach to examine “how three teams within one organisation used a new digital medium—Lotus Development Corporation’s Team Room technology” (Yates et al., 1997, p. 1).

In the analysis of the genres enacted by the teams within Team Room, the authors identified three distinct genre systems containing different genres:

- Meeting documentation: This genre system involves the communicative activity that “precedes and follows face-to-face meetings, that is, the announcement of the meeting’s logistics and agenda, and the subsequent distribution of minutes of the proceedings” (Yates et al., 1997, p. 3). Meeting documentation involves three genres (meeting logistics, meeting agenda, and meeting minutes). *Meeting logistics*, which was primarily identified by purpose and some common features of form, “is typically used to communicate the date, time, location, and duration of a planned meeting” (p. 3). The second genre in this system, *meeting agenda*, “is used to announce details of the purpose and content of the meeting” (Yates et al., 1997, p. 4). The researchers discovered that this genre has some common form features such as lists, text indentation, and highlighting. The third genre, the *meeting minutes*, used to “document the proceedings, decisions, and action items that occurred during the face-to-face meeting”(p. 4).
- Collaborative repository: The second genre system that was identified in the study is called “collaborative repository” and involves “using one document as a placeholder for later contributions to the database in the form of comments “nested” beneath the placeholder document” (Yates et al., 1997, p. 5). The teams use this genre system to support diverse activities such as coordinating schedules, initiating discussions, brainstorming, and creating topical information repositories similar to bulletin boards. The authors differentiate two genres existing in the collaborative repository: the placeholder and the response. The *placeholder* genre is identified by its purpose, whereas the *response* is identified by the combination of its relationship to the placeholder document and its purpose (“response”).
- Collaborative authoring: Collaborative authoring is a third genre system identified in teams' use of Team Room. Yates et al. (1997) differentiate three genres in this genre system: *circulated draft*, *reaction to draft*, and *final version*. Although the third genre was implicit it only rarely appeared in the Team Rooms.

According to the authors, “a key distinction of the genre systems [...] observed in the Team Room databases compared to their analogues in other electronic media is that the team members structured the genre systems themselves” (Yates et al., 1997, p. 8). This represents a valuable input for the development of the Open Knowledge Space in OPAALS, where a bottom up approach had already been chosen, assuming a more sustainable and participatory community when the individuals' needs and communication preferences are being recognised and implemented.

Genre systems

In an organisation, several genres are linked together forming so called *genre systems* (Yoshioka et al., 2001). The genres may also form a hierarchy, where a genre may be a specialisation of a more general genre ('is-a' relationship) or participate in the decomposition of a genre in several distinct

genres ('has-a' relationship) (Crowston & Williams, 2000). In this context, Yates & Orlikowski (1992) proposed the notion of sub-genre to express these hierarchical relations. In several research studies genre systems seem to be significant in organisation of the social, structural, temporal, and spatial dimensions of interaction generally, and powerful in structuring electronic interactions, where these dimensions of collaborative work have all shifted. The purpose of a genre is not an individual's private motive for communicating, but a purpose socially constructed and recognised by the relevant organisational community and invoked in typical situations (e.g., proposing a project, meeting to review project status). Spinuzzi (2000) considers two properties of genre systems: contingency and decentralisation. *Contingency* means organisational work is opportunistic sometimes and, thus, genre systems should not be considered formal specifications of coordinated activities (as workflow specifications) but rather under-specified communication patterns. *Decentralisation* means purpose is disseminated throughout a genre system and, thus, the whole combination of genres is necessary to explain an individual genre.

Orlikowski and Yates proposed six dimensions of genre systems as organising structures within a community: purpose (*why*), content (*what*), form (*how*), participants (*who/whom*), time (*when*) and place (*where*). They emphasise that these six dimensions are aspects of communicative interactions that are closely interconnected (Orlikowski & Yates, 1998):

- Purpose (*why*): The dimension of purpose means that a genre system delivers expectations and socially recognised functions of the systems and its components.
- Content (*what*): Besides the contents of the whole genre system and its individual components, this dimension provides expectations about the sequence of the elements.
- Form (*how*): This dimension includes several factors influencing the form of a genre system such as media, structuring devices, and linguistic elements.
- Participants (*who/whom*): An important factor of a genre system is the participants involved in a communicative interaction and their role.
- Time (*when*): While communication, participants often provide specific temporal expectations in a genre system (for example, deadlines).
- Place (*where*): This dimension provides location and place expectations not only for the individual genres but also for the entire genre system.

Orlikowski and Yates (1998) apply these characteristics of genre systems to three domains: research, design, and use. Within the scope of *research domain*, genre systems allows one to focus on communicative norms in practice. This standpoint is particularly useful and important in order to understand the nature of collaborative work mediated through technology (ibid). Applying genre system in *design domain* would help designers of collaborative technologies to make “explicit the

interdependence of a sequence of related genres in a genre system” and support “such linkage through technological means” (Orlikowski & Yates, 1998). Through the lens of genre systems in the design domain, the designers are able to add more general structures as well as providing tools to support users in building context- and domain-specific structures relevant to their collaborative activities, both existing and emergent. Applying genre systems to the “*use domain*” opens a new “user-friendly” dimension. Yates and Orlikowski point out that the developing of good tools is only one of the issues involved in the processes of effective collaboration (ibid). The interaction within an organisation is shaped by communicative norms, i.e. the genre systems.

Genre taxonomy

The *genre taxonomy* is often used in the context of communicative actions. It represents the elements of both genres and genre systems in terms of purpose, contents, participants, timing of use, place of communicative action, and form including media, structuring devices, and linguistic elements. Additionally, the genre taxonomy represents both widely recognised genres such as a report and specific genres such as a technical reports used in a specific company, because the difference between a widely recognised genre and a specific variant based on the more general genre sheds light on the context of genre use. Finally, it represents use and evolution of genre over time to help people to understand how a genre is relevant to a community where the genre is enacted and changed.

Yoshioka et al. (2001) propose a genre taxonomy as a knowledge repository of communicative structures or “typified actions” enacted by organisational members. The authors argue that the genre taxonomy may help “people to make sense and understand diverse types of communicative actions. The genre taxonomy is characterised by three features:

- The genre taxonomy represents the elements of both genres and genre systems, sequences of interrelated genres, as embedded in a social context considering their characteristics (Why, What, Who/Whom, When, Where, and How). This means that the genre taxonomy represents the elements in terms of their purpose, contents, participants, timing, place of communicative action, and form (Yoshioka et al., 2001).
- The difference between a widely recognised genre and a specific variant based on the more general genre sheds light on the context of genre use, i.e. the genre taxonomy represents both widely recognised genres such as a report and specific genres such as a technical reports used in a specific company
- The genre taxonomy represents use and evolution of genre over time.

Yoshioka et al. (2001) developed a prototype of a genre taxonomy by using the Process Handbook, a process knowledge repository introduced at the MIT. The authors applied the differentiation between several aspects of genre systems (why, how, who/whom, where, when) developed in Orlikowski & Yates (1998) to describe and analyse the genre taxonomy.

The initial purpose categories were established in the genre taxonomy based on speech act theory. Some categories were been modified and added using coding schemes which Yates and Orlikowski introduced in their empirical genre studies (Orlikowski & Yates, 1994; Yates et al., 1999; Yates et al., 1997). After category identification, Roger's thesaurus (Roger & Chapman, 1992) and WordNet (Fellbaum, 1998), an online lexical database for English, were used in order to clarify notions of each category. Several items were defined in this category: inform, commit, guide, request, express, decide, propose, respond, record and other. Since some genres, for example a memo, have multiple purposes, the genre taxonomy may help to differentiate between the primary and secondary purposes and to understand the relations between genre and social context (Yoshioka et al., 2001).

The prototype of the genre taxonomy currently contains both widely recognised genres and specific genres. Currently, the open set of widely recognised genres includes 14 genres: Business letter, memo, expense form, report, face-to-face meeting genre system, personal homepage, etc. Specific genres include the results of genre analysis of three studies undertaken by Yates and Orlikowski: Common Lisp project genres, Acorn project genres, and Team Room genre systems. As the authors pointed out, the genre taxonomy is implemented using the specialisation hierarchy, the decomposition hierarchy, flow dependency and the description field of an activity in the Process Handbook. The genre taxonomy described here has benefits as a knowledge repository through helping managers, consultants, and groupware designers to learn different communication methods and apply effective methods to their situational context and actions. The genre taxonomy provides diverse social contexts of communicative actions. It provides community members a source for new ideas related to the development of new communication methods, and to redesign existing communication methods. It also may help to resolve problems relating to communicative actions (Yoshioka et al., 2001).

Genre analysis

Genre analysis is often used in organisational communication research field. It helps to identify how an organisation structures its work. Organisations adopt and modify different genres to suit their coordination and collaboration needs (Yoshioka & Herman, 2000). A collection of genres is often called a *genre repertoire* (Yates et al., 1999; Orlikowski & Yates, 1994). Genre analysis is defined as “an approach to study communication patterns” (Antunes et al., 2006, p. 5). Antunes et al. (2006) apply genre analysis to the study of electronic meeting systems. One of the goals of genre analysis is to understand how virtual communities use digital communication to collaborate. This

knowledge is fundamental for informing the design of information systems. There are several problems the communities may face by using electronic meeting systems (Antunes et al., 2006, p. 5):

- “reduced organisational integration, neglecting many contextual cues and explaining factors necessary to make meeting outcomes usable within the organisation”;
- “lack of support to specific communities of users, stressing the dependency on a facilitator to configure and manage the technology”; and
- “lack of support to meeting occurrences that span long time periods”.

Antunes et al. (2006) suggest to apply genre analysis to the design of electronic meeting systems. The starting point of the genre analysis process is the analysis of communicative actions observed in an organisation and materialised through various media, for example memos, reports, e-mails, meetings, telephone calls, etc. (ibid). The authors suggest to see communicative actions as members of genres that have recognisable form, contents, and purpose. In addition, these genres may also be interpreted as being consistently related to each other, and thus aggregated in genre systems.

Yoshioka et al. (2001) suggest that one possibility of collecting and analysing communicative actions is to gather the following metadata: why (purpose), what (content), who (participants), when (timing), where (location), and how (form) (Yates & Orlikowski, 2002, Yoshioka et al., 2001). This can be done using several techniques, for example, document analysis, interviews, group discussions, ethnography, etc. However, the process is identified as complex for two reasons: Community ownership of genres must be guaranteed, and the requirement of analysing a large number of communicative actions.

(Electronic) meetings are defined in this study as twofold: they are a form of communication within organisations and they are genre systems. Communication, in turn, is considered as a fundamental component of organisations and serves the purposes of the information exchange and the transmission of meaning (Katz & Kahn, 1978).

To draw a conclusion, we refer to genre as a general concept including genre repertoire, system, taxonomy and genre analysis, and its role in epistemic cultures (communities and organisations) in general and the OPAALS research community and epistemic cultures existing within it in particular: The concept of genre offers researchers a powerful way of understanding work practices and interaction/communication norms and media usage. Examining the dynamics of genre concept development in communities could provide the explanations to the nature of the community and the social action it engages in (Miller, 1984).

2.3) Summary

So far, we introduced the concepts of discourse and epistemic culture including different approaches and directions of research. The concepts and studies chosen present a complementary view on communities as it is also discussed in work package 10, deliverable 10.5, where the focus is put on on-line communities and their practical implementation. Specifically the elaborations on genre systems concerning media usage and communication patterns emphasise the need for a systematic approach to the practical implementation of a joint on-line knowledge platform (the Open Knowledge Space). The next chapter will therefore provide an introduction to empirical studies, as a sound empirical analysis of discourse organisation models and platforms in and for OPAALS will offer valuable insights for Digital Ecosystems. Chapter 2 also provides linkage to work packages 7 and 8. It enriches, for example, studies of the Open Source phenomenon by a different focus regarding knowledge transfer and production in inter- and intra-organisational settings. An interesting question regarding this field of investigation is whether Open Source communities can be analysed by means of methodologies from epistemic cultures, as the socio-spatial relationships need to be addressed with different variables in on-line communities and – above this – in communities with members of different organisations and identities.

This chapter therefore reflects the richness of this field of investigation, regarding possible research approaches, and its interdisciplinary nature. As scientific knowledge becomes one of the major factors of cultural development, scholars have recognised the need and importance of the linking disciplinary fields to provide a comprehensive view of epistemic cultures and to facilitate application of knowledge in a specific area. This change of the research lens has influenced the growing interest within epistemic cultures in producing new knowledge through research that combines the skills and perspectives of multiple disciplines. From the socio-methodological point of view, the research focus moved from the “war between qualitative and quantitative paradigms” to recognising the convenience of mixing/combining their methodological approaches. Nowadays we have to handle a more global world, social reality and we are trying to realise more global understanding and knowledge production in this way. In the next section, the deeper view on the methodological background of discourse organisation in epistemic cultures including theoretical and practical implications will be presented and discussed.

3) Paradigms of epistemic cultures

3.1) Introduction

In an interdisciplinary, intercultural, semi-virtual and geographically dispersed community like OPAALS a variety of components needs to be considered to enable the understanding of the processes of communication, collaboration, knowledge production and dissemination. The OPAALS community is a mix of individuals and organisational units, each of them being shaped by different backgrounds: Country of origin, mother tongue, foreign language proficiency, formation, academical background, age, sex, culture, motivations, and interests, etc. As we discussed in the previous chapters, discourse and knowledge production are socially determined and each member of an epistemic culture represents an individual epistemic subject. On the one hand the process of knowledge and discourse production in an epistemic community depends on the epistemic subjects, on the other hand on the domain specific theoretical paradigms, research practice and genre systems.

In this chapter we examine the *production of discourse in the scientific community* with special focus on the differing research characteristics of the domains of natural science, social science and computer science.

When we use the term domain in context of epistemic communities we refer to the assumptions and methodologies agreed in the scientific community. For interdisciplinary work in general and within the OPAALS project in particular, it is very important to have a common language that enables the process of knowledge production as a result of planning, realising and discussing research and development together. We will therefore first examine the *theoretical paradigms* of the scientific domains which provide the basis for a coherent planning and realisation of the research process regarding knowledge and discourse production. In a second step we will review studies which analyse the actual practice applied in social science focused on the tendency to realise research by combining, mixing and integrating methods.

Finally we will analyse what a mixed method research design implicates and how it can be applied to the overall research framework of OPAALS (i.e. research in the field of digital ecosystems) and research regarding the OPAALS community development.

What do we mean by *domains of science*? A common understanding of what is meant by *science* refers to any systematic methodology which attempts to collect accurate information about the shared reality and to model this in a way which can be used to make reliable, concrete and quantitative predictions about events, past, present, and future, in line with observations.

Fields of science are commonly classified along two major lines: the *natural sciences*, which study natural phenomena, and the *social sciences*, which study human behaviour and societies. A principal

characteristic of these groups is that they are empirical sciences, which means the (produced) knowledge must be based on observable phenomena and capable of being tested for its validity by other researchers working under the same conditions.

Formal science, e.g. mathematics and logic, statistics and *computer science* is sometimes classified as the third group of science, having both similarities and differences with the natural and social sciences. It is similar to other disciplines in that it involves an objective, careful and systematic study of an area of knowledge; it is different because of its method of verifying knowledge, using a priori rather than empirical methods.

In a some “simplified” way we will distinguish between two ways of knowledge production (1) *a priori* - knowledge is independent of experience like in mathematics and computer science which to is characterise a constructive research and (2) *a posteriori* - knowledge is dependent on experience and based in both natural and social science on empirical and exploratory research.

Seipel and Riecker mention that the natural sciences like physics, biology and chemistry, and social sciences like economy, psychology, educational science and sociology are different regarding its object of research but not in the methodological process. This is the reason for a common or at least similar approach / paradigm in natural science and social science (deductive) (Seipel & Riecker, 2003).

The research process as a process of knowledge production in natural and social sciences is based on experience which can be a result of different theoretical paradigms and research practices (research design, data collection, and analysis) (Creswell, 1994; Seipel & Rieker, 2003; Bryman, 2004). OPAALS as an interdisciplinary research community has its representatives in three scientific domains: social science, computer science, and natural science.

In the case of scientific domains we have to analyse *research paradigms* as the common shared theoretical approaches and the research design and methods as the more practical implementations of the community in the process of knowledge and discourse production. According to Johnson and Onwuegbuzie (2004), the research paradigm is “...a set of beliefs, values, and assumptions that a community of researchers has in common regarding the nature and conduct of research. The beliefs include, but are not limited to, ontological beliefs, epistemological beliefs, axiological beliefs, aesthetic beliefs, and methodological beliefs“ (p.24). The term research paradigm represents an important part of the common aspects of the scientific domain besides the research design and methods (and genre systems) in the process of knowledge production.

3.1.1) Paradigms

Traditionally, the differentiation between qualitative and quantitative research (methodologies) is characterised by the belonging to distinctively different paradigm. The distinction between the paradigms relates to a number of levels concerning the production of knowledge and the research process: The rarefied level of epistemology, the level of “middle range” theory as elucidated in the theoretical framework, and the level of methods and techniques (Tashakkori & Teddlie, 1998). There is assumed to be a correspondence between epistemology, theory, and method. However, the distinction is most commonly applied at the methodological level, that is to say the process of data collection and the form in which the data is recorded and analysed. Although many researchers see themselves as belonging to one or the other paradigm, from the socio-methodological point of view the research focus nowadays is moving from the “war between qualitative and quantitative paradigms” to acknowledging the convenience of mixing/combining their methodological approaches. The boundaries of the paradigms are not clearly defined, neither in the theoretical framework nor in research practice. In this chapter we will highlight the traditionally defined “pure” paradigm and then discuss recent tendencies towards mixed methods research.

There is a tendency among some researchers (Tashakkori & Teddlie, 1998; Bryman, 2004) to treat epistemology and method as being synonymous. However, the epistemological assumptions do not dictate the specific kind of data collection and analytical methods researchers must use. In fact the assumptions of the scientific domain are guiding the design of the research process but the methodologies are not bound to an epistemic assumption. Creswell (1994) provides a more holistic view of a paradigm. The paradigms in the humanities and social science help us to understand different phenomena: their advanced assumptions about the social world, how science should be conducted, and what constitutes legitimate problems, solutions, and criteria of “proof”. We will follow the assumptions of the two paradigms given by Creswell in his book “Research Design, Qualitative & Quantitative Approaches” (1994) which reflects the classical distinction of qualitative and quantitative paradigm within a knowledge production framework.

3.1.2) Epistemological assumptions

To delimit the most important epistemological considerations of the main doctrines is a challenging undertaking because of a lack of a clear consent regarding borderlines or demarcation lines. Furthermore, neither the qualitative nor the quantitative paradigm follows exclusively one of the philosophical doctrines. We will provide a review of the streams positivism, realism, and interpretivism for the primary allocation of the two main paradigms of social sciences.

A particularly central issue in the process of knowledge production is the question whether the social world can and should be studied according to the same principles, procedures, and ethos as

can be found in natural science. Positivism affirms the importance of imitating the natural science and postulates an application of the methods of natural sciences to the study of social reality and beyond. Bryman (2004) outlines the following assumptions of positivism:

- “Only phenomena and hence knowledge confirmed by the senses can genuinely be warranted as knowledge (the principle of phenomenalism).
- The purpose of the theory is to generate hypothesis that can be tested and that will thereby allow explanations, laws to be assessed (the principle of deductivism).
- Knowledge is arrived at through the gathering of facts that provide the basis for law (the principle of inductivism).
- Science must (and presumable can) be conducted in a way that is value free (that is, objective).
- There is a clear distinction between scientific statements and normative statements and a belief that the former are the true domains of the scientists. This last principle is implied by first because the truth or otherwise of normative statements cannot be confirmed by the sense” (Bryman, 2004, p.11).

According to Johnson and Onwuegbuzie (2004), the quantitative purists articulate assumptions that are consistent with the positivist philosophy, they maintain that social science inquiry should be objective. A *quantitative study*, consistent with the quantitative paradigm, is an inquiry into a social or human problem, based on testing a theory, composed of variables measured with numbers and analysed with statistical procedures, in order to determine whether the predictive generalisations of the theory hold true. The researcher considers the reality as an objective independent of the researcher. In the context of the epistemological question, the quantitative paradigm proposes the distance and independence to the object being researched, for the objectivity of the study, “... thus in surveys and experiments, researchers attempt to control for bias, select as systematic sample, and be “objective” in assessing a situation” (Creswell, 1994, p.3). This means that the researchers' values are kept out of the study.

The qualitative purists (also called *constructivists* and *interpretivists*) reject what they call positivism. *Interpretivism* is a term that usually denotes an alternative to the position of the positivist orthodoxy that prevailed for several decades. It is predicated upon the view that a strategy that respected the differences between people and the objects of the natural sciences is essential and therefore requires the social scientist to grasp the subjective meaning of the social action. The interpretivism's intellectual heritage includes: Weber's notion of “*Verstehen*”; the hermeneutic-phenomenological tradition; and symbolic interactionism. These purists contend that multiple-constructed realities abound, that time- and context-free generalisations are neither desirable nor possible, that research is value-bound, that it is impossible to differentiate fully causes and effects,

that logic flows from specific to general (e.g., explanations are generated inductively from the data), and that knower and known cannot be separated because the subjective knower is the only source of reality (Bryman, 2004; Creswell, 1994; Johnson & Onwuegbuzie, 2004).

For a *qualitative researcher*, the only reality is the one constructed by the individuals involved in the research context. The researcher tries to minimise the distance between herself and those being researched. In the context of the epistemological question, there is interaction in qualitative research with those that are studied, “whether this interaction assumes the form of living with or observing informants over a prolonged period of time, or actual collaboration” (Creswell, 1994, p.6). According to the axiological issue of the role of values, the qualitative researcher admits the value-laden nature of the study and actively reports her values and biases as well as the value of the nature of the information gathered from the field.

The comparative summary of the paradigm which are purposed by the qualitative and quantitative purists are presented below (Table 4).

Table 4. Qualitative and Quantitative Paradigm Assumptions (Creswell, 1994)

Quantitative	Qualitative
Ontological assumption: <i>What is the nature of reality?</i>	
Reality is objective and singular, apart from the researcher.	Reality is subjective and multiple as seen by participants in a study.
Epistemological assumption: <i>What is the relationship of the researcher to the research?</i>	
The researcher is independent from that being researched.	The researcher interacts with what being researched.
Axiological Assumption: <i>What is the role of values?</i>	
Value-free and unbiased.	Value-laden and biased.
Rhetorical assumption: <i>What is the language of research?</i>	
<ul style="list-style-type: none"> - Formal - Based on set definitions - Impersonal voice - Use of accepted quantitative words 	<ul style="list-style-type: none"> - Informal - Evolving decisions - Personal voices - Accepted qualitative words
Methodological assumption: <i>What is the process of research?</i>	
<ul style="list-style-type: none"> - Deductive process - Cause and effect - Statistic design-categories isolated before study - Context-free 	<ul style="list-style-type: none"> - Inductive process - Mutual simultaneous shaping of factors - Emerging design-categories identified during research process

- | | |
|--|---|
| <ul style="list-style-type: none">- Generalizations leading to prediction, explanation, and understanding- Accurate and reliable through validity and reliability | <ul style="list-style-type: none">- Context-bound- Patterns, theories development for understanding- Accurate and reliable through verification |
|--|---|

This overview given by Creswell provides a background to the assumptions of the qualitative and quantitative paradigms. Nevertheless, the author points out that these are contrasts at the heuristic level; in the course of current research practice this ideal orientation is not represented by the studies conducted. Furthermore, there are many alternative models for research designs (Creswell, 1994).

3.1.3) Research Design

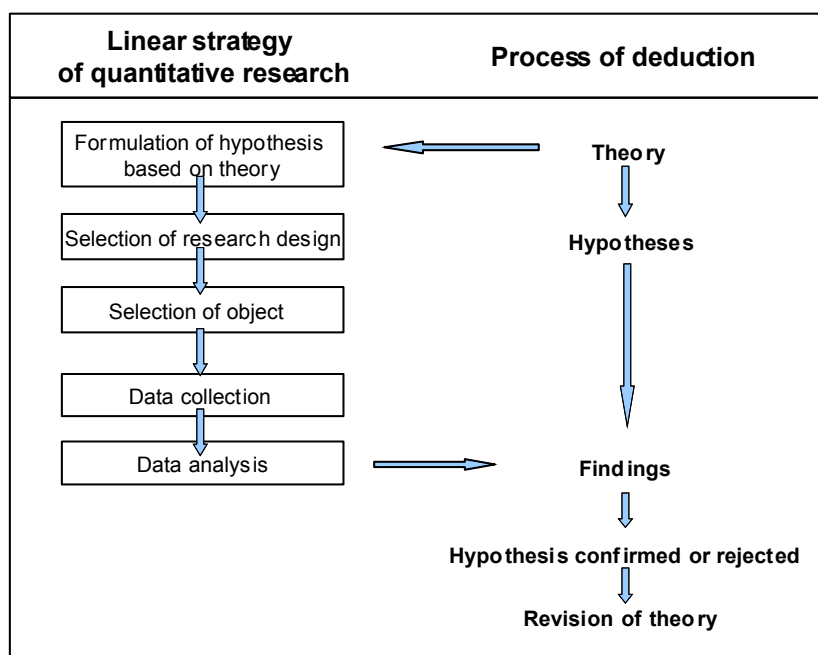
In the context of albeit drastic contrasts between qualitative and quantitative assumptions it seems that one precludes the other. Hence, purpose and typical research processes of the two paradigms are the next points we are going to discuss.

Quantitative research follows a deductive theory which represents the commonest view of the relationship between theory and social research, the *empirical research*. The research deduces a hypothesis (or hypotheses) that must be subjected to empirical scrutiny. The social scientist needs to both skilfully deduce a hypothesis and then translate it into operational terms (Bryman, 2004).

One also establishes a "cause effect" in the positivists methodology. A survey design provides a quantitative or numerical description of some fraction of the population – the sample. Through data collection, in turn, a researcher is enabled to generalise the findings from a sample of responses to the population (Creswell, 1994).

Qualitative research is concerned primarily with the process, rather than the outcomes or products. Qualitative researchers are interested in meaning and how people make sense of their lives, experiences, and their structures of the world. Qualitative research is the primary instrument for data collection and analysis, where data are not primarily mediated through inventories, questionnaires, or machines but rather by means of fieldwork, for example. This means that the researcher physically goes to the people, settings, sites, or institutions to observe or record behaviour in their natural environment. Qualitative research is descriptive as the researcher is interested in processes, meaning, and understanding gained through words or pictures. Furthermore, the process of qualitative research is inductive in that the researcher builds abstractions, concepts, hypotheses, and theories from details.

Figure 3. Process of deduction in linear quantitative research design (based on Seipel & Rieker, 2003; Creswell, 1994)

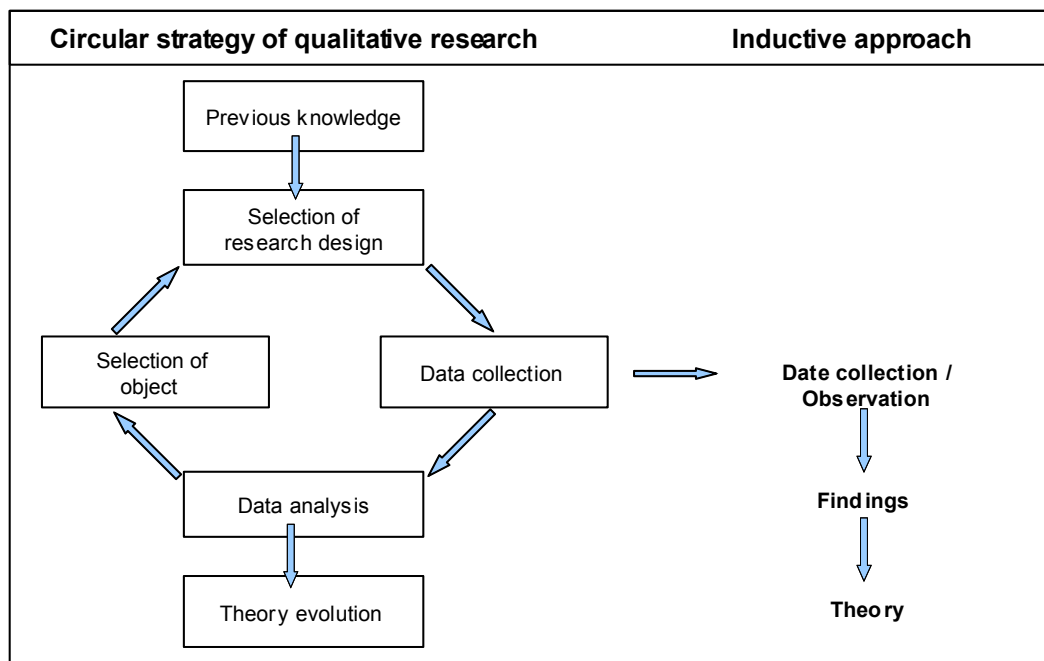


The character of a qualitative study is *exploratory*, one does not begin with a theory to test or verify. Instead, consistent with the inductive model of thinking, a theory may emerge during the data collection and analysis phase of the survey or be used relatively late in the research process as a basis for comparison with other theory (Creswell, 1994).

Qualitative research is a flexible, circular or iterative-cyclical process. With the inductive stance, theory is the outcome of this research approach. In other words, the process of induction involves drawing generalisable inferences out of observations (Seipel & Rieker, 2003).

In Figures 3 and 4 we illustrate the research process and the relationships between theory and research based on both paradigms. The quantitative paradigm is used for proving an existing theory whereas the purpose of the qualitative approach is to generate a theory. It seems that the theoretical division of the two paradigms referring to the theoretical and philosophical assumptions and the research purpose and design have traditionally accepted consent which do not allow to combine qualitative and quantitative research approaches or methods. Although some social science researchers perceive qualitative and quantitative approaches as incompatible, others believe that the skilled researcher can successfully combine them.

Figure 4. Circular strategy of qualitative inductive research design (based on Seipel & Rieker, 2003; Bryman, 2004)



3.1.4) Research methods

In a very simplified definition we understand the research method as a technique for data collection, which can involve a specific instrument, such as a questionnaire, structured interview, schedule, or participant observation, etc (Bryman, 2004).

In the literature, the methodological implementation of quantitative research design distinguishes between two types of *quantitative methods*: Experiments and surveys. Experiments are methods with the random assignment of subjects to treatment conditions and quasi experiments that use non-randomised designs. Included within experiments are single-subject designs. An experiment tests cause-and-effect relationships in which the researcher randomly assigns subjects to groups. Surveys include cross-sectional and longitudinal studies using questionnaires, structured interviews or self-completion questionnaires in order to collect a body of *quantitative or quantifiable data* in connection with two or more variables which are then examined to detect patterns of relationship between variables. In quantitative research we aim to generalise from a sample to a population. (Creswell, 1994; Bryman 2004).

Some of the studies represented and discussed in the previous chapters related to discourse organisation and epistemic cultures use quantitative methods of research: Scott and Timmerman

(2005; see chapter 2.2.3) developed a survey questionnaire in order to explore the use of workplace communication technologies. Besides measures of communication, computer, and writing apprehension the survey questionnaire results include the frequency report items for a set of technologies, and demographics information. In the latter sampling procedure, items to measure CMC apprehension and two additional media that has been gained their importance in organisational communication (instant messaging and online chat) have been added to the survey design. Smeby (2006; see chapter 2.2.1) uses a questionnaire approach in order to explore the extent to which and how beginning professionals of three different epistemic cultures (nurses, teachers and social workers) develop their knowledge and their assessment of the importance of different knowledge sources.

In the humanities and social sciences there are different traditions that correspond to different methods and types of *qualitative* data collection: ethnography and participant observation, focus groups, language analysis and grounded theory, interviewing, case studies, etc.

In the context of the action research, Marshak and Heracleous (2005; see chapter 1.3.4) apply the real time observation method in organisational meetings in order to analyse discourse from a situated symbolic action perspective. Cuel and Ferrario (2006; see chapter 2.2.3) suggest an active participation of the observer inside a firm, organisation, or epistemic community that is complemented by a series of qualitative interviews. Antunes et al. (2006; see chapter 2.2.4) suggest a genre analysis approach to study communication patterns which can be applied to the specific content of meetings. The described process contains the analysis of communicative actions by using qualitative observation methods. The goal of participating in several meetings as observer serves the purpose of developing the list of requirements for the design and implementation of electronic meeting systems.

All these methods can be used for collecting qualitative data correspondingly to the inductive research process proposed by the qualitative paradigm. Additionally, typical “qualitative methods” can be used in a more deductive way to produce quantifiable data; respectively, “quantitative” methods can be used in a more inductive way.

However, just as the deduction entails an element of inductive process, it is likely to entail a modicum of deduction. Once the phase of theoretical reflection on a set of data has been carried out, the researcher may want to collect further data in order to establish the conditions in which a theory will or will not hold. Such a general strategy is often called iterative, and involves a weaving back and forth between the data and theory. It is particularly evident in grounded theory (Bryman, 2004). We will point out this idea in the following discussion of interdisciplinary research and the use of a mixed method design.

Grounded theory researchers attempt to derive a theory by using multiple stages of data collection and the refinement and interrelationship of categories of information. Two primary characteristics of

this design are the constant comparison of data with emerging categories, and theoretical sampling of different groups to maximise the similarities and the differences of information (Krotz, 2005).

There is a rich score of research practice realising more complex research designs, data collection, and analysis. Often it is not possible to attribute them clearly as a paradigm (i.e. Action research). In the research practice of computer science and the natural science, mixed method research is not that uncommon. From the constructive and outcome oriented point of view we need different methodologies.

3.2) Why to Combine? - Theoretical and Methodological Implications

In current theoretical discussions and to a great extend in research practices, a mixed methods approach, i.e. combined and integrated research are already realised. There are different approximations to mixed methods and research design, their advantages and disadvantages, practical, rational and theoretical justification of the application.

By definition, mixed methods is a procedure for collecting, analysing, “mixing” or integrating both, quantitative and qualitative data, at some stage of the research process within a single study for the purpose of gaining a better understanding of the research problem (Tashakkori & Teddlie, 1998; Creswell, 2005). Johnson and Onwuegbuzie (2004) contend that epistemological and methodological pluralism should be promoted in educational research so that researchers are informed about epistemological and methodological possibilities and, ultimately, so that they are able to conduct more effective research. The rationale for mixing both kinds of data within one study is grounded in the fact that neither quantitative nor qualitative methods are sufficient by themselves to capture the trends and details of a situation. When used in combination, quantitative and qualitative methods complement each other and allow for a more robust analysis, taking advantage of the strengths of each approach (Tashakkori & Teddlie, 1998).

The philosophical, theoretical and methodological assumptions of the quantitative and qualitative paradigms provide orientation and the base to determine the paradigms of mixed methods research design. Nowadays, the discussion of philosophical orientation and paradigms of mixed method research is in progress. In the follow we will outline the philosophical, theoretical and methodological assumptions introduced by the current represents of the mixed method research from the point of view of the social science.

Tashakkori and Teddlie (1998) conducted a chronological review from logical positivism, post positivism to constructivism, pragmatism, and compatibility thesis. The authors describe pragmatism and the compatibility thesis which has given five convincing practical reasons for “coexistence” between two methodologies and their underlining paradigms:

- Both paradigms have, in fact, been used for years.
- Many evaluators and researchers have urged using both paradigms.
- Funding agencies have supported both paradigms.
- Both paradigms have influenced policy.
- So much has been taught by both paradigms (Tashakkori & Teddlie, 1998).

Tashakkori and Teddlie (1998) provide an evolutionary view of the methodological approaches in the social and behavioural sciences and describe three general stages in the evolution of mixed model studies;

- (a) Firstly, the acceptance of the use of mixed methods,
- (b) Secondly, the application of the distinctions that emerged during the paradigm wars to all phases of the research process, and
- (c) finally during the past 30 years a prosperous discussion of mixed method research.

In education and evaluation research, authors present the compatibility thesis based on a different paradigm, which some have called pragmatism.

3.3) Pragmatism

“We contend that epistemological and methodological pluralism should be promoted in educational research so that researchers are informed about epistemological and methodological possibilities and, ultimately, so that we are able to conduct more effective research” (Johnson & Onwuegbuzie, 2004, p. 15).

Both qualitative and quantitative methods may be used appropriately within any research paradigm. Questions of method choice are secondary to the questions of paradigm, which we define as the basic belief system or world view that guides the investigation, not only in choice of method but in ontologically and epistemologically fundamental ways (Tashakkori & Teddlie, 1998).

Tashakkori and Teddlie specify the following assumptions of pragmatism (Table 5):

Table 5. Paradigm of pragmatism (Based on Tashakorri & Teddlie, 1998)

Pragmatism	
Methods	Qualitative and Quantitative

Logic	Deductive and inductive
Epistemology	Objective and subjective points of view
Axiology	Values play a large role in interpreting results
Ontology	Accept external reality. Choose explanations that best produce desired outcomes
Causal linkages	There may be causal relationships, but we will never be able to pin them down

Pragmatism implies that values play an important role in conducting research and in drawing conclusions from their studies. The pragmatic researcher decides just in the study to pick and choose how and what to research and what to do. However, what does it mean for the research practice?

3.3.1) Rational for mixed method research

In his study “Integrating quantitative and qualitative research: how is it done?”, Bryman (2006) investigates how quantitative and qualitative approaches are integrated and combined in practice. The article is based on the content analysis of 232 social science articles in which the two paradigm were combined and used as a mixed method research approach. The author detected an interesting relation between the *rationales* for employing mixed-methods research approaches and the way it is put into practice (how the research framework itself is defined). He outlines dominating forms of mixed research method and designs. His analysis is based on the scheme of rationales, purposes for realising mix method research which was introduced by Greene et al. (1989). This scheme isolates five purposes for combining quantitative and qualitative research:

1. **Triangulation:** Convergence, corroboration, correspondence or results from different methods. In coding triangulation, the emphasis was placed on seeking corroboration between quantitative and qualitative data.
2. **Complementarity:** “Seeks elaboration, enhancement, illustration, clarification of the results from one method with the results from another” (Greene et al., 1989).
3. **Development:** “Seeks to use the results from one method to help develop or inform the other method, where development is broadly construed to include sampling and implementation, as well as measurement decisions” (Greene et al., 1989).
4. **Initiation:** “Seeks the discovery of paradox and contradiction, new perspectives of [sic] frameworks, the recasting of questions or results from one method with questions or results from the other method” (Greene et al., 1989).

- 5. Expansion:** “Seeks to extend the breadth and range of enquiry by using different methods for different inquiry components” (Greene et al., 1989).²

The result of the analysis states a discrepancy between the planning and the realised application of the mixed methods. Bryman (2006) points out that researchers even through planning a method mix in their research frameworks do not realise this in the research practice. Otherwise the researcher does not plan rationally the mix of methods in the research frameworks but in practice use mixed methods without specifying this as a rational.

Bryman (2006) based his analysis on three stages of the research process and analysed if the present studies do consider mixed methods (1) as a rational in the stage of formulation and planning of the research, (2) in the data collection and (3) data analysis.

The finding of his study suggest that there is relatively often a mismatch between the rationale for the combined use of quantitative and qualitative research and how it is used in practice (data collection and analysis). For several of the rationales, there is no evidence from the articles that quantitative and qualitative research approaches are combined in the way that the rationale would lead one to suspect. This is not always the case. In particular, when ‘instrument development’ and ‘sampling’ are the rationales, they are nearly always used in this way. Only one article was found claiming ‘instrument development’ as a rationale and only one article claiming ‘sampling’ as a rationale, but they did not report combining quantitative and qualitative research in these ways.

“An examination of the research methods and research designs employed suggests that on the quantitative side structured interview and questionnaire research within a cross-sectional design tends to predominate, while on the qualitative side the semi-structured interview within a cross-sectional design tends to predominate” (Bryman, 2006, p. 97).

The author points out the possibility that the reasons why multi-strategy research is undertaken have not been sufficiently articulated in the methodological literature, resulting in a lack of certainty about its uses. The second explanation is somewhat less negative in its connotations and implies that multi-strategy research frequently brings more to researchers’ understanding than they anticipate at the outset (Bryman, 2006).

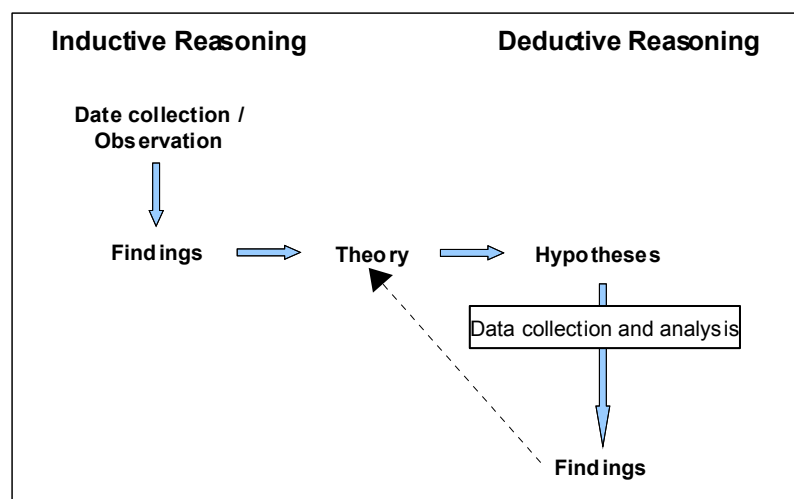
3.3.2) Mixed method research design

Tashakkori and Teddlie (1998) describe a research cycle / cycle of scientific methodology which fits into the pragmatism paradigm. Figure 5 illustrates the *methodological research design in the pragmatic paradigm*. Research projects concerning any substantive area of inquiry travel through

² The authors use a more detailed scheme of rationales, but in question of complexity we only allude to the simple scheme which represent the intention of his discourse.

this cycle at least once before it ends. Research projects may start at any point in the cycle: Some projects start from theories or abstract generalisations, while other start from observations. In many research reports, there is an initial attempt to inductively build a conceptual (theoretical) framework on the basis of previous findings. The obtained theoretical framework is then used as a basis for planning the course of the research. This may be in a chronological way or occur simultaneously as it is often accepted by pragmatism. Tashakkori and Teddlie (1998) classify this as a mixed model study with multiple applications within different phases of the study. Tashakkori and Teddlie (1998) define mixed model studies as studies that are products of the pragmatic paradigm and that combine the qualitative and quantitative approaches within different phases of the research process.

Figure 5. Methodological research design in the pragmatic paradigm (based on Tashakkori & Teddlie, 1998)



There may be single applications within phases of the study, such as quantitative (experimental) design, followed by qualitative data collection, followed by quantitative analysis after the data are converted. In this application the data would be converted to numbers using the “quantitising” technique (Tashakkori & Teddlie, 1998).

Generally there are three types of studies in social science: The mono method studies, mixed method studies and mixed model studies (Tashakkori & Teddlie, 1998). The traditional way of research was the mono method study conducted exclusively within one of the predominant paradigms, either quantitative or qualitative methodologies. Mixed method studies are those that combine qualitative and quantitative approaches into the research methodology of a single study or multi-phased study.

The authors use the following distinction:

- Sequential studies in which the researcher proposes to conduct a qualitative phase of the study and a separate quantitative phase of the study.
- Parallel/simultaneous studies in which the researcher proposes to conduct a qualitative phase and quantitative phase of the study parallel/simultaneously.
- Equivalent status design; the researcher conducts the study using both qualitative and quantitative approaches equally to understand the phenomenon under study.
- Dominant and less dominant studies; the researcher presents the study within a single, dominant paradigm with one small component of the overall study drawn from the alternative paradigm (e.g. quantitative study with a small qualitative interview, qualitative observation with limited number of informants, followed by a quantitative survey of a sample from a population).
- Design with multilevel use approaches; represents the highest degree of mixing paradigms. The researcher would mix aspects of the qualitative and quantitative paradigm at all or many methodological steps in the design (Creswell, 1995; Tashakkori & Teddlie, 1998).

Mixed model studies are defined as “studies that are products of the pragmatist paradigm and that combine the quantitative and qualitative approaches within different phases of the research process” (Tashakkori & Teddlie, 1998). The researcher should keep her assumptions explicit throughout all stages.

Johnson and Onwuegbuzie (2004) consider two important variables in the planning of mixed-model designs; firstly by mixing qualitative and quantitative approaches within and across the *stages* of research (in a simplified view, one can consider a single study as having three *stages*: stating the research objective, collecting the data, and analysing/interpreting the data). Secondly, one also may consider the dimension of paradigm emphasis (deciding whether to give the quantitative and qualitative components of a mixed study equal status or to give one paradigm the predominant status). *Time* ordering of the qualitative and quantitative phases is another important dimension, and the phases can be carried out sequentially or concurrently.

Qualitative research is often depicted as a research strategy whose emphasis on a relatively open-ended approach to the research process frequently produces surprises, changes of direction and new insights. However, quantitative research is by no means a mechanical application of neutral tools that results in no new insights. In quantitative data analysis, the imaginative application of techniques can result new understandings.

Johnson and Onwuegbuzie propose the “Mixed research process model” with seven-stages which is a conceptualisation of the mixed methods data analysing process. The seven stages of data analysis

are: data reduction, data display, data transformation, data correlation, data consolidation, data comparison, and data integration, represented as rectangles in Figure 6.

In a mixed method research the first step of mixed data analysing is *data reduction*, which involves reducing the dimensionality of the qualitative data and quantitative data, followed by the *data display*, the visualisation of the qualitative data and quantitative data. Once the data is visualised it is possible to interpret data in a parallel process. But when we aim to integrate the quantitative and qualitative data we have to conduct the *data transformation*, wherein quantitative data are converted into narrative data that can be analysed qualitatively and/or qualitative data are converted into numerical codes (Tashakkori & Teddlie (1998) specified this as qualitisied and quantitised data).

Data transformation is followed by *data correlation* and *data consolidation*, wherein both quantitative and qualitative data are combined to create new or consolidated variables or data sets. In the stage of *data comparison* the researches team compare the data from the qualitative and quantitative data sources. In the last stage of the mixed data analysis, *data integration*, both quantitative and qualitative data will be integrated into either a coherent whole or two separate sets of coherent wholes. This process permits to interpret both the qualitative and quantitative data collected in the mixed research process in a integrated way (Johnson & Onwuegbuzie, 2004).

3.3.3) Philosophical debates

Johnson and Onwuegbuzie (2004) assert that the philosophical debates in around the mixed method research will not end as a result of pragmatism, and nor that they should not end. Nonetheless, the authors agree with other authors involved in the mixed methods research:

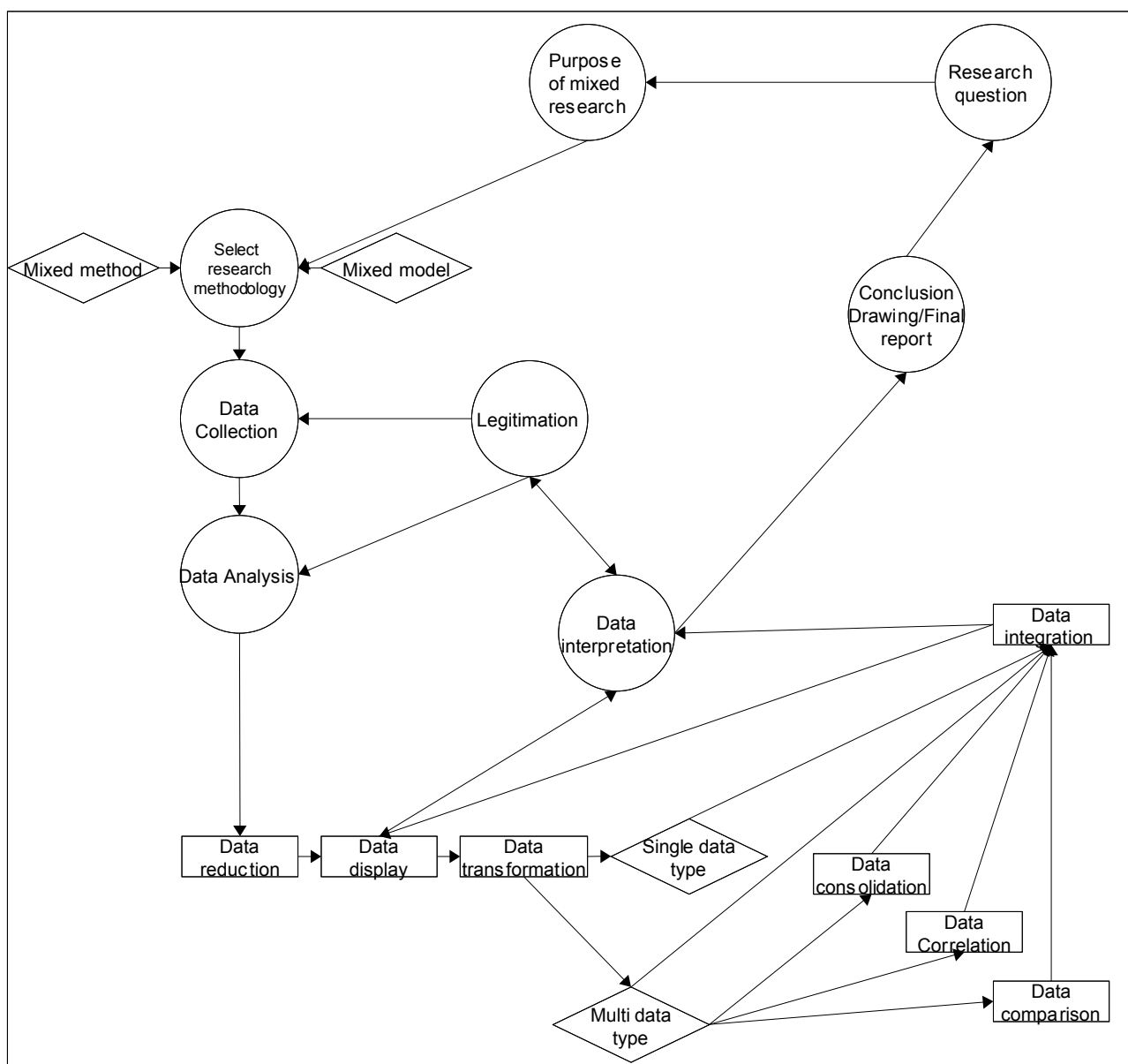
“... that consideration and discussion of pragmatism by research methodologists and empirical researchers will be productive because it offers an immediate and useful middle position philosophically and methodologically; it offers a practical and outcome-oriented method of inquiry that is based on action and leads, iteratively, to further action and the elimination of doubt; and it offers a method for selecting methodological mixes that can help researchers better answer many of their research questions” (Johnson & Onwuegbuzie, 2004, p.17).

In his article “*Notes on Pragmatism and Scientific Realism*”, Cherryholmes (1992) outlines the basic tenets of pragmatism, distinguishing it from the closely related philosophy of scientific realism. Pragmatists reject the idea that we can really know the truth about the world, and thus choose or dismiss theories based on their practical results rather than how "true" they are. The pragmatic researcher explicitly acknowledges that she has a desire for a certain kind of community based on her values, aesthetics, politics, and social and normative preferences, and chooses what and how to research based on them. “Not everything that works is desirable, not every belief that is

'true' is to be acted upon” (Cherryholmes, 1992, p. 14). Theories or strategies that violate the values and preferences of the researcher are ignored. Research literature is important not because it contains the truth but because it suggests constraints and opportunities (Cherryholmes, 1992).

Figure 6. Mixed research process model

Note: Circles represent steps (1–8) in the mixed research process; rectangles represent steps in the mixed data analysis process; diamonds represent components (Johnson & Onwuegbuzie, 2004).



The article “*Pragmatist philosophy and action research, Readings and conversation with Richard Rorty*” (Reason, 2003) explores Rorty’s pragmatism philosophy in relation to action research.

Reason (2003) discusses five basic characteristics of action research: (1) practical knowing, (2) democracy and participation; (3) ways of knowing; (4) human and ecological flourishing; and (5) emergent form, and identifies important links between the paradigm and philosophy of pragmatism contend by Rorty (see also Reason & Bradbury, 2001). Rorty's opposition toward the dualisms which haunts western philosophy is briefly described, his nonfoundationalist, anti-metaphysical pragmatics and his views on the contingency of the language that we use. Reason points out "... we can neither appeal to universal reason nor to an external reality as foundations for our claims." (Reason, 2003, p.107) so he consider that in a philosophical level we must move through a process of redescription. As another approximation to Rorty's position Reason represents the position of many action researches who agree that they cannot regard truth as a goal of theirs to inquire. Furthermore, there is the common aim to bring consensus. This means that an inquire should achieve coordination of behaviour.

As an characteristic of philosophical pragmatist the authors specify the assumption that ideas and practices should be judged in terms of their usefulness, workability and practicality and that these are the criteria of their truth, rightness and value, such it is a principle of action research to produce practical knowledge.

So in its ideal form, Rorty's culture of liberalism would have 'no room for the notion that there are non human forces to which human beings should be responsible', including the idea of a truth outside human imagination. Based on the assumption of the culture of liberalism by Rorty, Reason assumed that "... inquiry does not naturally converge on a consensus, to some end point of Truth or Reality or Goodness. Rather our only useful notions of 'true' and 'real' and 'good' are extrapolations from human created practices and beliefs, which will necessarily change over time" (Reason, 2003, p.106). In the 'non-foundational' view of Rorty "... there are no foundations for knowledge outside human discourse, no appeals to an ultimate Reality that can be made. It is a position that also questions the idea that human inquiry and science itself depends on a particular methodology." (Reason, 2003, p.106) The action researchers who would claim that the fundamental strategy of action research is to 'open communicative space' and help the emergence of 'communities of inquiry' just as the principle of action research of democracy and participation.

In the Rorty's point of view *language* is seen as making our world rather than representing the world:

"Truth cannot be out there – cannot exist independently of the human mind – because sentences cannot so exist, or be out there. The world is out there, but descriptions of the world are not. Only descriptions of the world can be true or false. The world on its own – unaided by the describing activities of human beings – cannot" (Rorty, 1989, p. 5).

The stimulating quality of Rorty's thought suggests that action researchers must find new language to describe their work, rather than be caught in the old academic metaphors of research. It is argued

that just as Rorty is re-describing philosophy, action researchers are re-describing inquiry “...action research is an orientation to inquiry rather than a methodology” (Reason, 2003, p.6).

3.4) Summary

Several studies discussed in the previous chapters combine qualitative and quantitative methods in order to gain a better view of the field and more complex information. A majority of the studies dealing with discourse and language uses statistical methods with linguistic parsers as well as qualitative analysis of the collected data. Auger (2005; see chapter 1.3.4) uses terminology extraction from electronic texts and followed identification of the semantic relations. Hilbert et al. (2006; see chapter 1.3.4) developed a complex architecture including knowledge sources (Discourse Marker Lexicon, Domain Ontology and text type structure ontology), auxiliary analysis components (Lexical and Morphological), and the discourse parser. In the context of the genres in organisational communication, Orlikowski and Yates (1994; see chapter 2.2.4) use mixed methods. They analyse 2000 transcripts of electronic mail messages both qualitatively and quantitatively. Genre analysis requires qualitative textual analysis of messages to understand the situations within which certain genres are invoked and their shared purpose, substance, and form. From a quantitative point of view, the textual analysis provides the basis for devising a coding scheme and for interpreting the patterns and trends identified. Background information and perspectives for interpreting these messages came from a series of two-phase, semi-structured interviews that were conducted after preliminary content analysis. The similar approach has been applied in the study of collaborative genres and genre systems in digital media (Yates et al., 1997; see chapter 2.2.4). The research design was based on the analysis of 492 messages posted in the Team Room databases over a seven-month period. Similar to the previous study, the authors developed a coding scheme for analysing messages based on the two dimensions of genre: form and social purpose. Form categories reflect the formatting features and linguistics practices used in the messages. Purpose categories reflect the socially recognisable purposes of messages interpreted both from the message content and the communication type designated by the message author (ibid). In addition to textual data analysis, interviews were conducted with members of each team. The qualitative interviewing has been applied before and after textual analysis and helped to ground and refine the interpretation of the genre systems the authors had identified and furnished insight into how and why they were used.

As in digital ecosystems and especially in this Network of Excellence we are dealing with a complex set of variables regarding communication/discourse and communities and knowledge, a mixed methods approach offers the possibility to analyse the variables and processes from different points of view. Moreover, this chapter also an introduction to the wide field of research methodologies and paradigms in order to firstly offer a sound point of departure for discussions regarding research designs (discussion that are unavoidable in interdisciplinary teams, being on the

one hand fruitful and on the other hand can tend to become a bottle-neck for the work progress), and secondly to use existing research designs and empirical studies outcomes for a complementary view on OPAALS and the planning of further work to be conducted in this project.

4) Sociolinguistic Framework

4.1) Introduction

Based on a selection of the aforementioned theories and applications, we developed a multidimensional sociolinguistic framework which provides the platform of our study and research design. The purpose of our study is to analyse the complex features of epistemic cultures existing in the Network of Excellence project community. In this research we combine qualitative and quantitative paradigms that help to provide different views of the same event, features or processes. From a qualitative point of view, our study relies on a systematic generation of statements from data that contains both inductive and deductive thinking. One goal is to formulate hypotheses based on conceptual ideas, and generate them by constantly comparing conceptualised data on different levels of abstraction. Another goal is to discover the participants' main concerns regarding their work in OPAALS and how they continually try to resolve it.

The theoretical background of our research design is based on the combination of different approaches and field studies in epistemic cultures, organisational communication, discourse theory, natural language processing, information retrieval, text mining, data mining and knowledge discovery, sociolinguistics, computer linguistics, social network, social science etc.

There are several factors affecting the success and effectiveness of community building and development. However, the most important factor for the success of any community is its people. It is therefore crucial to understand people's behaviour in such communities (Rafaeli et al., 2004). The major advantage of the applying the Socio-linguistic Framework analysis to the epistemic cultures existing in the OPAALS community is the ability to see its members in two different situations:

1. As developers and designers,
2. and as users.

In this context, the number and complexity of factors to be analysed are increasing drastically. How do developers produce, share knowledge? How does the “culture of organisation” affect knowledge production?

By means of the sociolinguistic framework, we seek to understand in depth how knowledge moves within networks, how social aspects affect knowledge management, and the role of domain-specific language and communication. The framework design was developed in the context of the empirical study of discourse organisation in epistemic cultures within the project.

The study is designed as a long-term study and the first on-line questionnaire has already been sent out and data gathered. Furthermore, additional data has also been gathered during meetings in form of notes, recordings, and photographs. The first results of the the long-term study will be published

in form of an article and are also being made available for work packages 5 and 10 which integrate tasks regarding the development of the OPAALS Open Knowledge Space.

4.2) Framework Domains: Multidimensionality of Analysis

Four dimensions are taken into account in the sociolinguistic framework: Language, Knowledge, Communication, and Network. The co-evolution of these factors, shown in Figure 7, forms a comprehensive image of the needs or conditions for the community development. A domain-specific focus of language analysis adds a number of advantages to our research study. It is perhaps obvious that the problem of the resolution of semantic ambiguity becomes more manageable once the focus is narrowed to the language used in a specific domain (Farghaly & Hedin, 2003).

Since the OPAALS community contains researchers not only with different cultural/geographical, but also educational backgrounds, it is important to analyse the specifics and dynamics of language usage in context of knowledge representation, sharing, and intercultural communication. At the interpersonal level, the opportunity of having the full gamut of semantic representations introduces a more complex possibility of sharing experience through communication and this process enriches individual experience with possible new meanings in a circular and recursive process, allowing the social construction of cultural models in an essential dynamic between creativity and sharing (Proietti, 2003).

4.2.1) Language

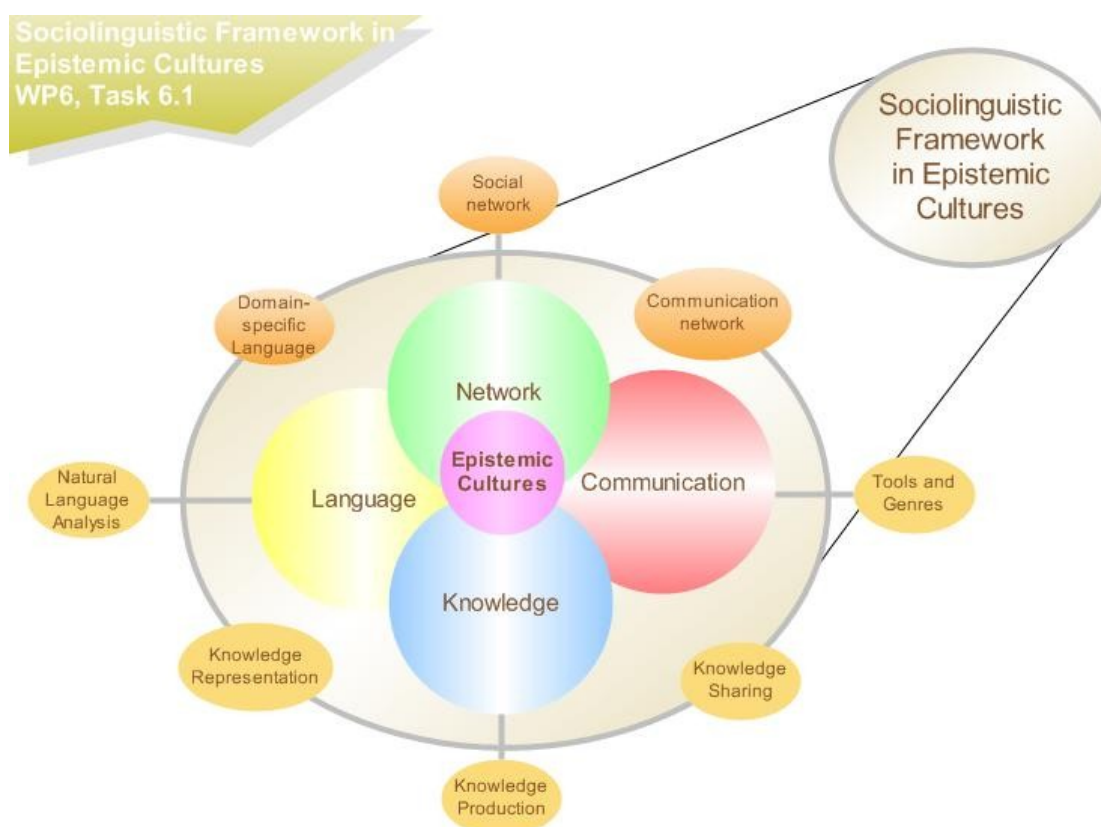
The complexity of NLP systems stems from the intricacy of language itself (Farghaly & Hedin, 2003). In context of knowledge production and movement within and among communities, it is important to understand how linguistic items are selected and organised to serve the specific cognitive aims of a particular knowledge domain. Based on text documents in form of attachments, publications, OPAALS web site and emails, the socio-linguistic framework provides a comprehensive linguistic study design of language specifications and characteristics such as word choice, metaphors usage, sentence features etc. based on the sentence and word level.

4.2.2) Knowledge

One of the most important goals of community building, especially in context of OPAALS, is knowledge production, representation and sharing. The community plays a critical role in determining what we take to be true, and how we develop knowledge. Knowledge is always situated in a place, time, conditions, practices and understandings (Van House, 2004).

According to Constant (1984), technology should be seen as a form of knowledge. Similarly, Knorr-Cetina (1999) has proposed that scientific practices can be understood as “epistemic cultures” that bind together tools, knowledge, and specific knowledge production mechanisms. From this point of view, the sociolinguistic framework can be seen as a multidimensional interacting system that not only provides a comprehensive analysis but also helps to build and organise communities and their relations and interactions with others. This might also be connected to the correlation between knowledge production and communication goals. Applying the sociolinguistic framework, we aim to find dependencies between the concrete selection of the applied knowledge representation tools and the functional-communicative intention of the community members. Furthermore, knowledge exists in a social context, and is created in epistemic cultures. Knowledge transfer is the process through which one network member is affected by the experience of another (Inkpen & Tsang, 2005).

Figure 7. Sociolinguistic Framework



4.2.3) Communication

Communicative competence is best viewed not as knowledge of a static system of “rules“ but as an open and dynamic potential that can always be enhanced (de Beaugrande, 1997). Communication, defined generally as transfer of information or resources, is common among socially related people (Hady et al., 2005).

Information and Communication Technology (ICT) in turn has changed science and work life during the last years. It facilitates collaboration in dispersed research teams, provides new publication platforms (open access, online-journals, web logs, and databases) and enriches interpersonal communication with new tools. In short, ICT rapidly pervades our daily lives regarding both private and work life. It seems that ICT is adopted in different ways by different epistemic cultures (Matzat, 2002; Nentwich, 2004). For example, computer scientists have created techniques for knowledge creation (RFCs, Open Source), which may be a first hint that computer scientists have a different discourse organisation compared to other epistemic cultures. Email is said to be the “killer-application” of the internet (Ducheneaut & Bellotti, 2001), based on empirical evidence that emailing is used for communication, collaboration (document management and exchange) and personal information management (Ducheneaut & Bellotti, 2001). However, there are more tools than email to facilitate scientific work, and communication and collaboration traditions differ according to the disciplinary domain. Hence, these differences influence the internet usage of each researcher (Matzat, 2002). The question is: To what extend are there differences? One further assumption is that working and communication traditions/practices and the personal interests of researchers influence the usage and acceptance of computer-mediated-communication-tools.

4.2.4) Network

We identify two types of the network: Communication network and social network. Communication network analysis provides an avenue for analysing and comparing information flows in epistemic cultures. In combination with social representation of organisations, we are able to compare information flows with formally defined work processes. We aim to find particularly interesting results in terms of communicative practices regarding the different scientific domains of the test community.

Another advantage is the possibility to connect collaborations on formal languages (computer languages) with natural language practices: How far does the formal and logic framework of computer languages affect the communicative behaviour of its users when communicating within their own domain and with another domain?

A social network then describes a group of social entities and the pattern of inter-relations among them (Hady et al., 2005). The distributed character of the OPAALS community and splitting into individuals, groups (task teams), institutions and work package levels introduce a complex social structure of the overall organisation. Additionally, as geographically distributed community, the project links people across countries and organisational units. Several factors (distance, organisational affiliation, cultural differences, etc.) make the process of community building and evolution more difficult. Cultural differences can easily lead to communication difficulties and to misinterpretation. Language also introduces a very basic barrier of communication (Wenger et al., 2002).

5) Conclusion

This deliverable aimed at discussing the two overarching concepts of OPAALS: Language and community. By choosing an overview and introductory approach, it also intended to

- foster the understanding of a social science framework within the OPAALS community, and
- to initiate interdisciplinary discours(es).

By depicting and discussing relevant research frameworks for OPAALS, it has become evident that the project community is embedded in a highly complex 'ecosystem'. Figure 8 draws a map of the project as a 'living organism', which refers to the intrinsic dynamic nature of natural language (on which the project is built), and to the overall digital ecosystem metaphor.

Figure 8. OPAALS' living organism

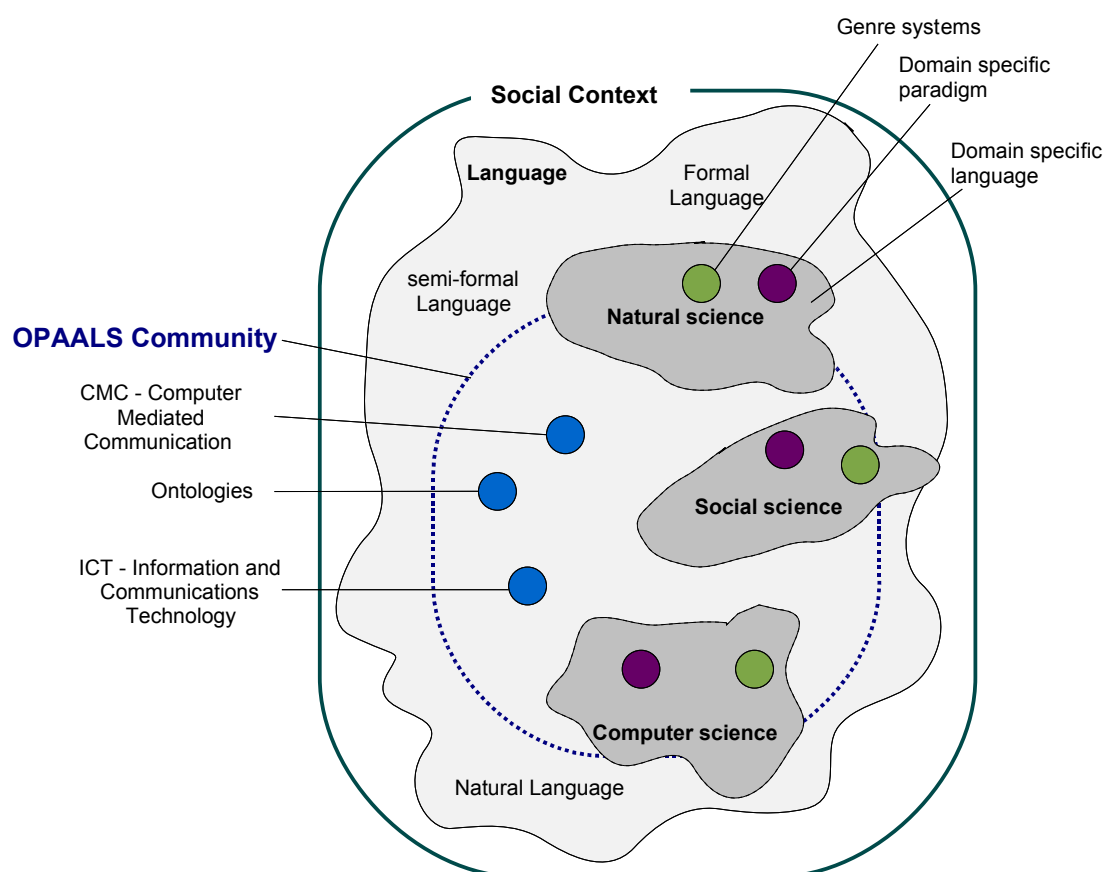


Figure 8 also reveals the highly complex nature of OPAALS, the different domains which are dependent on varying communicative frameworks, and after all on the specific context.

Through using exploratory and empirical research designs, the social science domain may supply initial ideas, concepts, models, data to support the construction process and development carried out by the computer science domain which contains impulses from the natural science knowledge repository regarding biological phenomena. Furthermore, the social science domain is able to observe the effects of the developed solutions as well as to accompany the process of design and implementation in the computer science domain. Finally, it provides the analysis of the outcome (final product) including usability, importance, acceptance by the communities and their members, changes caused by the implementation process. Therefore, existing connections of social science related tasks and approaches to work packages from computer science were therefore mentioned in this deliverable. Moreover, potential and already planned work based on this deliverable already were pointed out.

To sum up, one of the important issues we are still facing in OPAALS is clarifying the goals the community wants to achieve through the interaction of the three domains. With the following 'sub'-goals we therefore close our deliverable and open further discussions and research directions.

1. Effective interdisciplinary collaboration (“operational ecosystem”): Through the development and accompanied analysis of language, communicational norms and styles and community building, OPAALS aims to achieve an effective interdisciplinary collaboration as well as to collect set of requirements for further implementation
2. Developing a common language: The development of a “common language” is important not only for the successful communication and collaboration within the OPAALS community but also for design modelling including rules and policies.
3. Community's Philosophy and behavioural aspects: Besides language, communication, and collaboration, the overall “community's philosophy” and behavioural rules are important aspects in the organisational development. Each community (or organisation) has a particular amount of language codes and styles of spoken and written language as well as the possible interaction strategies within community.

6) Bibliography

- Allott, R. (2001). *The great mosaic eye: Language and evolution*. Sussex: Book Guild Ltd.
- Andersen, P. B. (2003). Semiotic models of algorithmic signs. In K.-H. Rödiger (Ed.), *Algorithmik - Kunst - Semiotik* (pp. 165-193). Heidelberg: Synchron Wissenschaftsverlag der Autoren.
- Antunes, P., Costa, C. J., & Pino, J. A. (2006). The use of genre analysis in the design of electronic meeting systems. *Information Research*, 11(3).
- Auger, A. (2005). *Applying natural language processing techniques to domain knowledge capture*. Paper presented at Intelligence Analysis, McLean, VA.
- Bamford, J., & Bondi, M. (Eds.). (2005). *Dialogue within discourse communities: Metadiscursive perspectives on academic genres*. Tübingen: Max Niemeyer Verlag.
- Baumann, K.-D. (2001). LSP text as a knowledge system. In F. Mayer (Ed.), *Languages for Special Purposes: Perspectives for the New Millenium* (Vol. 1, pp. 105—113). Göttingen: Gunter Narr.
- Baumard, P. (1999). *Tacit knowledge in organizations*. London: Sage.
- de Beaugrande, R. (1997). *New foundations for a science of text and discourse: Cognition, communication, and the freedom of access to knowlegde and society* (Vol. 61). Norwood, NJ: Ablex Publishing Corporation.
- Bergquist, M., & Ljungberg, F. (1998). *From memo to intranet: Genre-analysis as a way to understand change in organizational communication*. Paper presented at Proceedings of IRIS 21, Aalborg, Denmark.
- Boje, D., Oswick, C., & Ford, J. (2004). Language and organization: The doing of discourse. *Academy of Management Review*, 29(4), 571—577.
- Bourdieu, P. (2006). Language and symbolic power. In: A. Jaworski & N. Coupland (Eds.), *The discourse reader* (2nd ed., pp. 480-490). London: Routledge.
- Bouwan, H., & van de Wijngaert, L. (2002). Content and context: An exploration of the basic characteristics of information needs. *New Media Society*, 4(3), 329—353.
- Brown, J. S., & Duguid, P. (1991). Organizational learning and communities of practice: to-wards a unified view of working, learning and innovation. *Organization Science*, 2(1), 40-57.
- Bryman, A. (2004). *Social research methods* (2nd ed.) New York: Oxford University Press.
- Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research*, 6(1), 97—113.
- Campbell, J. (2006). Media richness, communication apprehension and participation in group videoconferencing. *Journal of Information, Information Technology, and Organizations*, 1, 87—96.

- Chalaby, J. K. (1996). Beyond the prison-house of language: Discourse as a sociological concept. *British Journal of Sociology*, 47, 684—698.
- Cherryholmes, C. H. (1992). Notes on pragmatism and scientific realism. *Educational Researcher*, 21(6), 13—17.
- Chklovski, T., Ratnakar, V., & Gil, Y. (2005) *User interfaces with semi-formal representations: A study of designing argumentation structures*. Paper presented at Proceedings of Conference on Intelligent User Interfaces, San Diego, CA.
- Constant, E. W. (1984). Communities and hierarchies: Structure in the practice of science and technology. In R. Laudan (Ed.), *The Nature of Technological Knowledge: Are Models of Scientific Change Relevant?* (pp. 27—46). Dordrecht: Reidel.
- Cowan, R., David, P., & Foray, D. (2000). The explicit economics of knowledge codification and tacitness. *Industrial and Corporate Change*, 9(2), 211—253.
- Creswell, J. W. (1994). *Research design: Qualitative & quantitative approaches*. Thousand Oaks, CA: Sage Publications.
- Cronin, B. (2003). Scholarly communication and epistemic cultures. *New Review of Academic Librarianship*, 9(1), 1—24.
- Crowston, K., & Williams, M. (2000). Reproduced and emergent genres of communication on the world-wide web. *The Information Society*, 16(3), 201—216.
- Cuel, R., & Ferrario, R. (2006). *The impact of technology in organizational communication processes: Toward constructivism*. Paper presented at EGOS 2006, Bergen, Norway.
- Daft, R. L., & Lengel, R. H. (1986). A proposed integration among organizational information requirements, media richness, and structural design. *Management Science*, 32, 554—571.
- Davenport, T. H., De Long, D. W., & Beers, M. C. (1998). Successful knowledge management projects. *MIT Sloan Management Review*, 39(2), 43—57.
- de Saussure, F. (2003). *Grundlagen der allgemeinen Sprachwissenschaft*. Berlin: DeGruyter.
- de Souza, C. S. (2005). *The semiotic engineering of human-computer interaction*. Cambridge, MA: MIT Press.
- Ducheneaut, N., & Bellotti, V. (2001). Email as habitat: An exploration of embedded personal information management. *Interactions*, 8(5), 30—38.
- Evers, H.-D. (2000). Globalization, local knowledge, and the growth of ignorance: The epistemic construction of reality. *Southeast Asian Journal of Social Science*, 28(1), 13—22.
- Evers, H.-D. (2005). Global knowledge: The epistemic culture of development. In R. Hassan (Ed.), *Local and Global: Social Transformation in Southeast Asia* (pp. 3-17). Leiden and Boston: Brill.
- Fairclough, N. (1995). *Critical discourse analysis*. Harlow: Longman.

- Fairclough, N. (2001). *Language and power* (2nd ed.). London: Longman.
- Fairclough, N. (2003). *Analysing discourse: textual analysis for social research*. London: Routledge.
- Farghaly, A., & Hedin, B. (2003). Domain analysis and representation. In A. Farghaly (Ed.), *Handbook for Language Engineers* (pp. 21-58). Stanford, CA: SCLI Publications.
- Fellbaum, C. (1998). *Wordnet: An electronic lexical database* MA: MIT Press, <http://www.cogsci.princeton.edu/~wn/w3wn.html>.
- Foucault, M. (1972). *The archaeology of knowledge*. New York: Pantheon.
- Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, NF 100, 25-50.
- Frege, G. (1960). On sense and reference. In P. Geach & M. Black (Eds.), *Translations from the philosophical writings of Gottlob Frege* (pp. 56-78). Oxford: Basil Blackwell.
- Georgakopoulou, A., & Goutsos, D. (1997). *Discourse analysis: An introduction*. Edinburgh: Edinburgh University Press.
- Giere, R. N. (2002). Discussion note: Distributed cognition in epistemic cultures. *Philosophy of Science*, 69, 637—644.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation design. *Educational Evaluation and Policy Analysis*, 11(3), 255—274.
- Grosz, B. J., & Sidner, C. L. (1986). Attentions, intentions, and the structure of discourse. *Computational Linguistics*, 12(3), 175-204.
- Haas, P. M. (1992). Introduction: Epistemic communities and international policy coordination. *International Organization*, 46(1), 1—35.
- Hady, W. L., Lim, E.-P., Tan, T. T., & Pang, H. H. (2005). Mining social network from spatio-temporal events. *Computational & Mathematical Organization Theory*, 11(2), 97—118.
- Herre, H., & Heller, B. (2006). Semantic foundations of medical information systems based on top-level ontologies. *Knowledge-Based Systems*, 19, 107—115.
- Herring, S. C. (2001). Computer-mediated discourse. In D. Schiffrin, D. Tannen & H. Hamilton (Eds.), *The Handbook of Discourse Analysis* (pp. 612—634). Oxford and Cambridge, MA: Blackwell.
- Hilbert, M., Lobin, H., Bärenfänger, M., Lungen, H., & Puskas, C. (2006). *A text-technological approach to automatic discourse analysis of complex texts*. Paper presented at Proceedings of KONVENS, Konstanz.
- Hjørland, B. (2002). Domain analysis in information science: Eleven approaches - traditional as well as innovative. *Journal of Documentation*, 58(4), 422—462.

- Howard, P. N. (2002). Network ethnography and the hypermedia organization: new media, new organizations, new methods. *New Media & Society*, 4(4), 550-574.
- Inkpen, A. C., & Tsang, E. W. K. (2005). Social capital, networks, and knowledge transfer. *Academy of Management Review*, 30(1), 146—165.
- Johnson, B. R., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14—26.
- Jurafsky, D., & J. H. Martin. (2000). *Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics*. Upper Saddle River, NJ: Prentice-Hall.
- Kamp, H. & Reyle, U. (1996). A Calculus for First Order Discourse Representation Structures. *Journal of Logic, Language and Information*, 5, 297–348.
- Kamp, H., van Genabith, J. & Reyle, U. (2007). Discourse Representation Theory: An updated survey. DRAFT of an article to appear in the new edition of the *Handbook of Philosophical Logic*. Retrieved from <http://www.ims.uni-stuttgart.de/~hans/hpl-drt.pdf>
- Katz, D., & Kahn, R. (1978). *The social psychology of organizations* (2nd ed.) New York, NY: John Wiley & Sons.
- Kay, M. (2004). Introduction. In: R. Mitkov (Ed.), *The Oxford handbook of computational linguistics*. Oxford: Oxford University Press.
- Knorr-Cetina, K. (1997). Sociality with objects: Social relations in postsocial knowledge. *Theory, Culture & Society*, 14, 1—30.
- Knorr-Cetina, K. (1999). *Epistemic cultures. how the sciences make knowledge*. London: Harvard University Press.
- Knorr-Cetina, K., & Bruegger, U. (2002). Traders' engagement with markets - a postsocial relationship. *Theory, Culture & Society*, 19, 161—185.
- Krotz, F. (2005). *Neue Theorien entwickeln. Eine Einführung in die grounded theory, die heuristische Sozialforschung und die Ethnographie anhand von Beispielen aus der Kommunikationsforschung*. Köln: Herbert von Halem Verlag.
- Lacey, N. (2000). *Narrative and genre: Key concepts in media studies*. New York, NY: St. Martin's Press.
- Marshak, R. J., & Heracleous, L. (2005). A discursive approach to organization development. *Action Research*, 3(1), 69—88.
- Matzat, U. (2002). Academic communication and internet discussion groups: What kind of benefits for whom? In B. Batinic, U.-D. Reips & M. Bosnjak (Eds.), *Online Social Sciences* (pp. 383—402). Göttingen: Hogrefe & Huber.
- McCarthy, E. D. (1996). *Knowledge as culture: The new sociology of knowledge*. London: Routledge.

- Miller, A. (2003). Philosophy of language. In: Shand, J. (Ed.), *Fundamentals of philosophy* (pp. 262-296). London, New York: Routledge.
- Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech*, 70, 151—167.
- Miller, K. (1995). *Organizational communication: Approaches and processes*. Belmont, CA: Wadsworth Publishing Company.
- Moore, J., & Wiemer-Hastings, P. (2003). Discourse in computational linguistics and artificial intelligence. In A. Graesser, M. Gernsbacher & S. Goldman (Eds.), *Handbook of discourse processes* (pp. 439–486). Mahwah, NJ: Erlbaum.
- Nentwich, M. (2004). Cyberscience, methodology, and research substance. In M. Consalvo, N. Baym, J. Hunsinger, K. Bruhn Jensen, J. Logie, M. Murero & L. Regan Shade (Eds.), *Internet Research Annual. Selected Papers from the Association of Internet Researchers Conference 2000-2002, Volume 1* (pp. 77—85). New York: Peter Lang.
- Orlikowski, W., & Yates, J. (1994). Genre repertoire: Norms and forms for work and interaction. *Administrative Science Quarterly*, 39, 541—574.
- Orlikowski, W., & Yates, J. (1998). *Genre systems: Structuring interaction through communicative norms* Sloan MIT WP 4030: CCS WP 2005.
- Parker, I. (1992). *Discourse dynamics*. London: Routledge.
- Phillips, N., & Hardy, C. (2002). *Discourse analysis: Investigating processes of social construction*. Thousand Oaks, CA: Sage Publications.
- Podolny, J., & Page, K. (1998). Network forms of organization. *Annual Review of Sociology*, 24, 57-76.
- Pogner, K.-H. (2005). Discourse communities and communities of practice: On the social context of text and knowledge production in the workplace. Working paper, <http://ep.lib.cbs.dk/download/ISBN/x656494750.pdf>.
- Proietti, R. (2003). Language, communication and development of the self. In C. B. Grant (Ed.), *Rethinking Communicative Interaction* (pp. 69—85). Amsterdam: John Benjamins Publishing Company.
- Putnam, L. L., Phillips, N., & Chapman, P. (1996). Metaphors of communication and organization. In S. R. Clegg, C. Hardy & W. Nord (Eds.), *Handbook of Organizational Studies* (pp.375-408). London: Sage.
- Putnam, L. L., & Fairhurst, G. T. (2001). Discourse analysis in organizations. In: L. L. Putnam & F. M. Jablin (Eds.), *The new handbook of organizational communication. Advances in theory, research, and methods* (pp. 78-136). London: Sage.
- Rafaeli, S., Ravid, G., & Soroka, V. (2004). *De-Lurking in Virtual Communities: A Social Communication Network Approach to Measuring the Effects of Social and Cultural Capital*. Paper presented at Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04), Hawaii.

- Reason, P., & Bradbury, H. (Eds.). (2001). *Handbook of action research: Participative inquiry and practice*. London: SAGE Publications.
- Reason, P. (2003). Pragmatist philosophy and action research: Readings and conversation with Richard Rorty. *Action Research*, 1(1), 103—123.
- Rice, R. E., & Williams, F. (1984). Theories old and new: The study of new media. In R. E. Rice (Ed.), *The new media: Communication, research and technology* (pp. 55–80). Beverly Hills: Sage Publications.
- Rice, R. E. (1993). Media appropriateness: Using social presence theory to compare traditional and new organizational media. *Human Communication Research*, 19(4), 451—484.
- Rindflesch, T. C., & Fiszman, M. (2003). The interaction of domain knowledge and linguistic structure in natural language processing: Interpreting hypernymic propositions in biomedical text. *Journal of Biomedical Informatics*, 36(6), 462—477.
- Roger, P. M., & Chapman, R. L. (1992). *Roget's international thesaurus*. New York, NY: Harper Collins.
- Rorty, R. (1989). *Contingency, irony, and solidarity*. Cambridge: Cambridge University Press.
- Santos, F. M., & Eisenhardt, K. M. (2005). Organizational boundaries and theories of organization. *Organization Science*, 16(5), 491-508.
- Schiffrin, D. (1994). *Approaches to discourse*. Oxford and Cambridge, MA: Blackwell.
- Schreibman, S., Siemens, R., & Unsworth, J. (Eds.). (2004). *A companion to digital humanities*. Oxford: Blackwell.
- Scott, C. R., & Timmerman, E. (2005). Relating computer, communication, and computer-mediated communication apprehensions to new communication technology use in the workplace. *Communication Research*, 32(6), 683—725.
- Seipel, C., & Rieker, P. (2003). *Integrative Sozialforschung: Konzepte und Methoden der qualitativen und quantitativen empirischen Forschung*. Weinheim & München: Juventa Verlag.
- Sheth, A., & Ramakrishnan, C. (2003). Semantic (web) technology in action: Ontology driven information systems for search, integration and analysis. *EEE Data Engineering Bulletin*, 26(4), 40—48.
- Smeby, J.-C. (2006). *Epistemic cultures among beginning professionals*. Paper presented at 23rd Nordic Conference of Sociology, Turku, Finland.
- Spinuzzi, C. (2000). *Grappling with distributed usability: A cultural-historical examination of documentation genres over four decades*. Paper presented at Proceedings of the 17th annual international conference on computer documentation, New Orleans.
- Sure, Y., Akkermans, H., Broekstra, J., Davies, J., Ding, Y., Duke, A., et al. (2003). On-To-Knowledge: Semantic Web Enabled Knowledge Management. In N. Zhong, J. Liu & Y. Yao (Eds.), *Web Intelligence* (pp. 277-300). Springer-Verlag.

- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage Publications.
- Van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & Society*, 4(2), 249-283.
- Van Dijk, T. A. (2000). *Ideology and discourse. A multidisciplinary introduction*. Retrieved February 20, 2007, from <http://www.discourses.org/Unpublished/Ideology%20and%20discourse.pdf>
- Van House, N. A. (2004). Epistemic machineries of environmental online communication. In A. Sharl (Ed.), *Environmental Online Communication* (pp. 199—208). London: Springer.
- Wahlster, W., & Kobsa, A. (Eds.). (1988). *Computational Linguistics. Special Issue on User Modeling*, 14(3), 101-103.
- Wahlster, W. (1991). User and discourse models for multimodal communication. In J.W. Sullivan & S. W. Tyler (Eds.), *Intelligent user interfaces* (pp. 45-67). New York: ACM Press.
- Webber, B. L. (2003). Computational perspectives on discourse and dialogue. In D. Schiffrin, D. Tannen & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 798-816). Oxford: Blackwell.
- Weick, K. E. (1979). *The social psychology of organizing*. Reading, MA: Addison-Wesley.
- Weick, K. E. (1995). *Sensemaking in organizations*. London: Sage Publications.
- Wenger, E., McDermott, R., & Snyder, W. M. (2002). *Cultivating communities of practice*. Boston, MA: Harvard Business School Press.
- Whitley, E. A. (1991). Two approaches to developing expert systems: A consideration of formal and semi-formal domains. *AI & Society*, 5, 110—127.
- Wodak, R., & Meyer, M. (Eds.). (2001). *Methods of critical discourse analysis*. London: Sage.
- Woodilla, J. (1998). Workplace conversations: The text of organizing. In D. Grant, T. Keenoy & C. Oswick (Eds.), *Discourse and organization* (pp. 31—50). Thousand Oaks, CA: Sage Publications.
- Yates, J., & Orlikowski, W. (1992). Genres of organizational communication: A structural approach to studying communication and media. *Academy of Management Review*, 17(2), 299—326.
- Yates, J., Orlikowski, W., & Rennecker, J. (1997). *Collaborative genres for collaboration: Genre systems in digital media*. Paper presented at the Thirtieth Annual Hawaii International Conference on System Sciences, Hawaii.
- Yates, J., Orlikowski, W., & Okamura, K. (1999). Explicit and implicit structuring of genres: Electronic communication in a Japanese R&D organization. *Organization Science*, 10, 83—103.
- Yates, J., & W. Orlikowski. (2002). Genre systems: Structuring interaction through communicative norms. *Journal of Business Communication*, 39(1), 13—35.

Yoshioka, T., & Herman, G. (2000). *Coordinating information using genres* CCS No. 214, Sloan No. 4127. Center for Coordination Science MIT.

Yoshioka, T., Herman, G., Yates, J., & Orlikowski, W. (2001). Genre taxonomy: A knowledge repository of communicative actions. *ACM Transactions on Information Systems*, 19(4), 431—456.