FRAMING UNIVERSITY SMALL GROUP TALK: KNOWLEDGE CONSTRUCTION THROUGH LEXICAL CONCEPTS

YUN PAN

Thesis submitted for the degree of Doctor of Philosophy

Integrated PhD in Educational and Applied Linguistics

Newcastle University

Faculty of Humanities and Social Sciences

School of Education, Communication and Language Sciences

November 2017

DECLARATION

I hereby certify that this thesis is based on my original work. All the quotations and citations have been duly acknowledged. I also declare that this thesis has not been previously or currently submitted for any other degree at Newcastle University or other institutions.

Name: Yun PAN

Signature:

Date: 22/11/2017

ABSTRACT

Knowledge construction in educational discourse continues to interest practitioners and researchers due to the conceptually "natural" connection between knowledge and learning for professional development. Frames have conceptual and practical advantages over other units of inquiry concerning meaning negotiation for knowledge construction. They are relatively stable data-structures representing prototypical situations retrieved from real world experiences, cover larger units of meaning beyond the immediate sequential mechanism at interaction, and have been inherently placed at the semantic-pragmatic interface for empirical observation. Framing in a particular context – university small group talk has been an under-researched field, while the relationship between talk and knowledge through collaborative work has been identified below/at the Higher Educational level. Involving higher level cognitive activities and distinct interactional patterns, university small group talk is worth close examination and systematic investigation.

This study applies Corpus Linguistics and Interactional Linguistics approaches to examine a subset of a one-million-word corpus of university small group talk at a UK university. Specifically, it provides a detailed examination of the participants' framing behaviours for knowledge construction through their talk of disciplinary lexical concepts. Analysis reveals how the participants draw upon schematized knowledge structures evoked by particular lexical choices and how they invoke expanded scenarios via pragmatic mappings in the ongoing interaction. Additionally, it is demonstrated how the framing moves are related to the structural uniqueness of university small group talk, the contextualized speaker roles and the institutional procedures and routines.

This study deepens the understanding of the relationship between linguistically constructed knowledge and the way interlocutors conceptualize the world through institutionalized collaboration, building upon the existing research on human reliance upon structures to interpret reality at both the conceptual and the action levels. The study also addresses interaction research in Higher Educational settings, by discussing how the cognitive-communicative duality of framing is sensitive to various contextual resources, distinct discourse structures and task procedures through the group dynamics.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr Adam Brandt and Dr Christopher Leyland. I am grateful to have had their guidance and support. Thanks too to Professor Steve Walsh, Dr Dawn Knight and Dr Rachelle Vessey for their academic guidance.

My thanks also go to the NUCASE research group in the School of Education, Communication and Language Sciences. Thanks to the group members Qi Chen, Kazuki Hata, Khadija El-Wakai, Cihat Atar and Michael Grez for their warm support and precious friendship. Thanks are due to Professor Paul Seedhouse, Dr Peter Sercombe and Professor Sue Robson for their constructive feedback in my panel sessions which has been a great help to refine my research design and evaluate the research findings. Special thanks are also due to Dr Muge Satar Coen and Dr Jane Evison for examining this thesis and for offering very constructive feedback and suggestions in the viva.

Huge thanks are due to the teaching/research staff at Lancaster University for holding the Summer School for Corpus Linguistics research during the academic year 2015/16. Thanks to Professor Tony McEnery and Professor Paul Baker for the interesting discussions on the Corpus Linguistics methodology through personal communication. Special thanks to Dr Paul Rayson for generously sacrificing his weekend time to help me deal with the data cleaning issues through the interface of Wmatrix.

I am grateful to my fellow students in the PGR research suite 2.41. Special thanks are due to Yuzhu Peng, Rui Zhang, Miaomiao Zuo and Yuqi Xiao for their trust, understanding and support to help me get through all the tough days. Thanks to all my good friends in China.

Final thanks go to my beloved family and particularly to my parents, Jing Ji and Hongjiang Pan, for their endless support and encouragement. I will never be able to express how lucky and proud I am to be their daughter. This piece of work is dedicated to them, as a tiny tribute to their sacrifice and love.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Setting the scene

Knowledge construction in educational discourse has never failed to catch researchers' eyes due to the conceptually "natural" connection between knowledge and learning. The learning practice, -formal or informal, conscious or unconscious, guided or unguided- to a large extent depends on individual learners' choices about what new ideas and perceptions are to be accepted into their established views of the world. The Piagetian (1972) perspective emphasizes the collective construction of knowledge through the interaction between individual's epistemic structures and his/her experience of the environment. This constructivist perspective, although is not necessarily the only way to view the process of knowledge acquisition, is useful to examine human reliance on structures of knowledge in their learning practice.

Knowledge empowers individuals and promotes greater participation in social affairs thus is the fundamental precondition and an enduring stimulus for scientific development, economic growth, social justice and political democracy. The value of knowledge has been so widely and deeply accepted that there is a consensus around the world on the crucial importance of being involved in educational activities. However, there seems to be a gap between what people really get from formal education and what they have expected from it. One cause of the problem may be that it is quite difficult to support continuous and active learning under most of the existing educational policies. A sustainable development agenda is thus needed based on a conceptual shift of what it really means to have knowledge. The basic perspective adopted in this study is that knowledge is more about the construction of social reality than the reception of informational input. The interaction between learners' cognitive development

and social actions plays a central role in defining what knowledge is and how it is to be known.

Progressively developed conceptions of knowledge in universities are associated with distinct study behaviours of learners when compared with what has been found in institutions below Higher Education. Ideas and perspectives are found to be formed and evaluated more critically, and beliefs and justifications are constructed and negotiated in a more relational, tentative, complex, and integrated way among adult learners. There is a cognitive development concerning how to understand the validity of the source, structure, transmission and stability of knowledge, which contributes to shaping university students' sense of control of knowledge and motivation for learning.

Among a broad range of institutional activities in university settings, collaborative work in small groups are of particular interest to researchers during the past decades. Knowledge is expected to be constructed based on shared interpretation as a result of mutual understanding pursued and maintained by participants in groups. Small group work, therefore, has been taken as a key component of academic endeavours to encourage active learner participation and to optimize pedagogical support. Knowledge construction in such context is viewed as a process of mediating individual cognition and expectation through institutional interaction.

This study pays primary attention to knowledge construction in the setting of *university small group talk*. University small group talk is one type of organized interaction taking place in Higher Education contexts. A typical university small group compromises 4 to 12 participants and is characterized by sustained interaction (see Walsh and O'Keeffe 2010). In this study, I

consider university small group talk as *knowledge construction discourse* (van Aalst 2009) in which participants' language use is a powerful tool for cognition and action to shape speaker roles, social relationships, and patterns of educational community practice. The spoken language use constituting participants' talk of discipline-specific lexical concepts is taken as a unique linguistic register to represent a very specialized way of thinking about the academic world at the Higher Educational level. *Lexical concepts* (Evans 2006) are viewed to provide access to conceptual knowledge structures at larger scales which constitute schematized units of meaning. The *negotiation of meaning* (Pica 1994) around target lexical concepts contributes to organized interpretation and patterned communication; thus, is positioned at the centre of knowledge construction.

The constructed knowledge is the main focus of this study. In this opening chapter, an overview of the study was given (Section 1.2) to introduce the central research themes. The research aims and research questions were articulated (Section 1.3). The chapter ended with an outline of the thesis structure (Section 1.4).

1.2 Research overview

This study explores the linguistic practice of participants of university small group talk in terms of their framing behaviours for knowledge construction. In this section, three central aspects of this study –knowledge construction, university small group talk, and frame and framing will be briefly introduced.

1.2.1 *Knowledge construction*

Knowledge is abstracted from human life experience oriented to a specific social class or community (Marx and Engels 1845). The construction of knowledge is, thus, dependent on meaning representation collaboratively achieved across social contexts (Peirce 1934; Clark and Brennan 1991; Hardin and Higgins 1996). In this study, the construction of knowledge is viewed as an aspect of meaning negotiation (Rommetveit 1992; Pica 1994; Fauconnier 1997a; Gass 1997; Stahl 2006) for knowing, a process in which language use, local situations, and social activities are shared among participants (Cheung 2009) within a mediational framework (Roschelle and Teasley 2012) for meaningful communication.

Knowledge construction is greatly facilitated by language use (Spuzic and Nouwens 2004) since the latter constitutes signifying systems (Teubert 2007) involving the translation of perceptions and thoughts and the transfer of knowledge from one individual to another (Gergen 1985). Language provides the basis for thinking, divides experience into categories, and gives it meaning (Gergen 1985; Burr 1995; Lantolf and Thorne 2000). Meaning negotiation is therefore taken as the basic social practice of a community to achieve knowledge construction in collaborative context (Basturkmen 2002; Cheung 2009).

Meaning negotiation relies on the combination of units of knowledge as mental representations of life experience. In the human conceptual system, a concept is represented by a conceptual structure which takes a form of a semantic structure when encoded in language (Hamawand 2016). Semantic structures constitute meanings associated with words which are referred to as *lexical concepts* (Evans 2006; 2009). Lexical concepts have semantic values which are open-ended and dependent on the context of utterance (Evans 2006; Clarks 1983; Coulson 2001; Evans and Green 2006). Words are thus believed to have *meaning potentials* (Fauconnier 1992; Allwood 2003) rather than absolute meanings calculated

independent from context. Meaning potentials of words provide access sites to rather than straightforward identifications of non-linguistic conceptual knowledge structures (Evans 2006; 2009). Therefore, it is the pragmatic selectivity rather than the semantic exclusiveness that is attributed to lexis in use (Sweetser and Fauconnier 1996). The construction of knowledge through language use is accordingly taken as a process of representing situated meanings projected by semantic potentials associated with linguistic forms for specific communicative purposes.

As meaning is associated with utterance as a function of language in use (Evans 2006), it is essential to understand knowledge construction through analysing discourse (Cole 1996; Wenger 1998; Engestrom 1999; Koschmann et al. 2000; Koteyko 2006). Structural configurations of discourse are at the centre of meaning production and negotiation with situated interpretations of linguistic forms mutualized among participants (Fauconnier 1997b). Discourse skills are thus regarded as one of the key competences to participate in educational activities (Uccelli et al. 2014; Heller and Morek 2015) with knowledge as the product of constructive discursive processes (Bereiter and Scardamalia 1989).

Previous research has addressed the relationship between knowledge construction and discourse structures (e.g. Bereiter and Scardanalis 1989, 1996; van Aalst 2009). The relationship, though can be perceived in various dimensions, has been most explicitly addressed by Bernstein (1999). Bernstein's work had been conducted in the context of pedagogical interaction (at the macro level), focusing on the social principles that regulate the classification of knowledge which is made and practiced in educational institutions. Based on such classification, he distinguished two modalities of discourse – *vertical discourse* and *horizontal discourse* with forms of knowledge as the criteria. He argued that the distinctions

between the two kinds of discourse in terms of their practice, distributive principles, social relations involved and acquisition of knowledge contribute to shaping different modes of learning and context management (see Section 2.2.2 for a more detailed account of Bernstein's work and Section 6.2.3 for an intensive discussion). The distinction provides a useful way of examining educational discourse from a structural perspective, which is compatible with the view adopted in the current study to approach knowledge construction in such context.

Bernstein's (1971) work has implications for a deeper understanding of both the content and the organization ("ways of knowing", p. 57) of knowledge constructed in educational discourse. Nevertheless, it suffers limitations in terms of its empirical validity (little first-hand findings from empirical studies), research scope (mainly conducted in institutions below Higher Education) and analytical focus (on products rather than processes of knowledge construction). From here, a gap in the research literature emerges. This study will address this gap by focusing on a discursive setting which is strikingly different from what was observed by Bernstein (see Section 1.2.2).

Following the call for a more comprehensive view of knowledge and cognition beyond the laboratory research setting (Hutchins 1995), scholars in the field of sociology suggest the great importance of social context in knowledge construction (Kittleson and Southerland 2004). In other words, knowledge is viewed as socially constructed (Vygotksy 1978; Hyland 1999; Roth 1993; Mercer 1995; Raymond 2003; Stahl 2004; Raymond and Heritage 2006; Arvaja 2007; Heritage 2013). Social interaction constituting moment-to-moment sequential management mediates participants' *relative epistemic positions* (Hayano 2011; Heritage 2011; 2012; Heritage and Raymond 2012) towards particular fields of knowledge. Participants'

engagement in the dynamic change of their relative epistemic positions contributes to the progressivity of talk and *epistemic progression* (Gardner 2007). Knowledge construction, accordingly, is treated as a structural layer of the organization of talk-in-interaction (Balaman and Sert 2017).

Specifically, Labov and Fanshel (1977) developed a five-fold classification of knowledge in relation to local states of interaction concerning how wide the gap is between interlocutors in terms of their possession of particular knowledge. Building upon the classification, Kamio (1997) introduced the concept of *territories of information* (Kamio 1997) which allowed the development of a fuller sociological perception of knowledge domains (Berger and Luckmann 1967; Lukes 1974; Abbott 1988). Heritage (2011; 2012; 2013) drew together the above ideas and coined the notions of *epistemic status* and *epistemic stance*: the former involves participants' joint recognition of relative access to particular knowledge and the latter concerns local expression of social relationships through the design of turns at talk. The link between knowledge construction and social interaction justify the focus on the sequential organization of communication.

Educational talk as a variant of social interaction involves an intensive use of academic language for knowledge construction (Heller and Morek 2015). Linguistic features typical of academic language represent different conceptions of what knowledge is, where it is from and how it is used. Such linguistically represented conceptions are used to define and differentiate disciplines (Groom 2007). Academic disciplines are regarded as distinct *epistemic domains* (Stivers and Rossano 2010) with an established relative epistemic status among participants and greater stability of what occurs as a consequence of such state of communicative affairs (Heritage 2013; see also Tannen and Wallat 1987).

Research in disciplinary discourse during the past decades seemed to embrace the assumption that different academic disciplines feature different knowledge structures and skills (e.g. Kolb 1981; Bridgeman and Carlson 1984; Casanave and Hubbard 1992; Hyland 1999). Among the studies conducted in this field, Becher and his colleagues' work (Becher 1987a, 1987b, 1989, 1994; Becher and Trowler 2001; Neumann et al. 2002) is the best-known to provide a disciplinary categorization based on the nature of knowledge and patterns of linguistic practice. He explicated the relationship between the two and positioned different disciplines within different knowledge domains. As a result, disciplinary features reflected in participants' language use can be systematically investigated, compared and evaluated with the comparative parameters controlled (see Section 2.2.3 for a more detailed account of Becher's work).

The categorization contributes to a better understanding of how knowledge structures and linguistic behaviours affect each other to characterize academic language use. Nevertheless, it risks incompleteness and oversimplification since it overlooks cross-disciplinary similarities in terms of language use for knowledge construction. In addition, little attention has been paid to knowledge construction in communication at the micro level (see Stivers et al. 2011; Levinson 2012). This study will address these two gaps by conducting close observations of talk-in-interaction and try to identify if, and how, knowledge is constructed in similar/different ways across academic disciplines.

1.2.2 *University small group talk*

The data analysis of this study is based on a one-million-word corpus of spoken academic discourse: the Newcastle University Corpus of Academic Spoken English (NUCASE) (Walsh 2014). The main data set constitutes audio recordings of university small group talk sessions across three academic disciplines: Applied Linguistics and Education, Bioinformatics and Marine Engineering. The types of interactions vary, including tutored and non-tutored student group meetings, seminars, student project meetings, consultancy meetings between students and company delegations, etc.

A relationship between the level of knowledge construction and the organization of collaborative work has been identified at the Higher Educational level (see Gash 2015; Atwood et al. 2010; Arvaja et al. 2007). Previous research has examined how collaborative talk among students affects their knowledge of subject content and practice procedure (e.g. Mason 1998; Syh-Jong 2007). Collaborative work in small groups is found to involve higher level cognitive activities thus contribute to higher level academic achievements (see Hillocks 1986; Johnson and Johnson 1999; Barron 2003; Amato and Amato 2005). Small group talk has, therefore, become a standard feature of academic interaction across disciplines (Westgate and Corden 1993).

In spite of the benefits of small group work to facilitate knowledge acquisition, challenges arise to understand how specific goals and conventions of the university contribute to shaping collaborative talk (Atwood et al. 2010). There is a significant change on the conceptualization of learning at the Higher Educational level, which has an impact on the relationship between participants involved in educational activities. The role of university teachers/tutors has been shifted from the knowledge authority to what is more like a facilitator, a counsellor, a working partner and a co-communicator. At the same time, university students' rising awareness of

independent reasoning, critical thinking (Cantwell and Andrews 2002), learning motivation (Boekaerts and Minnaert 2006), and even the sense of resistance to their conventionalized academic identity (Benwell and Stokoe 2002) has become increasingly relevant to understanding ways of communication in their learning practice. All of these would inevitably blur the boundary between academic authority and apprentice, contributing to a more equal power distribution for teaching and learning. This is the first aspect of the setting which renders it different from contexts which have been examined previously -the research focusing on the pedagogically oriented design of students' collaborative work (e.g. Howard 2001; Blatchford et al. 2003; Macdonald 2003; César and Santos 2006). It should be noted that the small group talk in NUCASE does have pedagogical goals when it is defined and approached in the context of teaching and learning, while how the goals are to be achieved are expected to be different from what is in traditional lecture settings. What makes NUCASE distinct is the fact that the pedagogy is more task-oriented than procedure-oriented, being highly context-sensitive and negotiable.

The second aspect of this setting concerns the relationship between its institutionality (Drew and Heritgage 1992) and knowledge construction. As students majoring a particular academic discipline in NUCASE share similar knowledge background and are at the same stage of study, their relative professional expertise and status (Heritage 2012; 2013) are known to each other. The case, however, does not necessarily apply to the sessions where teachers/tutors or company delegations are involved in the talk with students. In such sessions, while both parties may self-identify as knowledgeable, the negotiation of meaning for knowledge construction is more likely to be managed in a guided way. As one shall see in the data analysis of this study, the institutionally defined speaker role is closely related to one's responsibility and orientation to claim possession of particular knowledge and to determine

how it is to be represented and transmitted (see Kamio 1997; Maynard 2003; Stivers et al. 2011) in the group talk.

While previous studies have conducted research into the relationship between discourse structures and knowledge forms in educational contexts (e.g. Cazden 1986; Bereiter and Scardamalia 1989; Bernstein 1971, 1986, 1999; Berrill 1991; Coultard 1992; Richmond and Striley 1996; Koschmann et al. 2000; Basturkman 2002; van Aalst 2009), they mainly focused on the impacts of organizational, psychological, technological, or cultural factors on the group performance and atmosphere. A discursive approach has not been fully developed to collaborative talk in university settings in terms of how knowledge is constructed through the sequential formation of communicative actions. What has been addressed insufficiently is the discursive mechanism of processes of knowledge construction viewed from a micro perspective, although a handful of studies have already been conducted to fill the gap (e.g. Markee 1995; Stokoe 2000; Benwell and Stokoe 2002).

1.2.3 *Frame and framing*

In this study, the investigation of knowledge construction in university small group talk involves the investigation of units of meaning which represent valid structures of knowledge. As mentioned in Section 1.2.1, competent language users get access to non-linguistic knowledge (or encyclopedic knowledge) by linguistically encoding conceptual structures in their mind. Such a coherent conceptual knowledge structure is referred to as a *frame*, first defined by Minsky (1975) as "a data-structure for representing a stereotyped situation" (p. 3). The idea that human reasoning, perception and language use are expected to be interrelated within larger structures of knowledge was shared by many scholars in the 1970s and 80s (e.g.

Newell and Simon 1972; Norman 1972; Schank 1972; Abelson 1973; Schank and Abelson 1975; Lakoff 1977; Rumelhart 1980).

In a further extension of the idea, Fillmore (Fillmore 1961; 1968; 1969; 1971; 1975; 1978; 1982a; 1982b; 1985a; 1985b; 1986; 1987) developed the theory of Frame Semantics to interpret and apply the notion of *frame* in the research of empirical semantics. A *frame* is defined in the theory as "a system of categories structured in accordance with some motivating context" (Fillmore 1982a: 381). The connection between linguistic knowledge and encyclopedic knowledge (Cienki 2012) is viewed from a constructive perspective that frames constituting schematized experience are invoked by human language use in a system of communication. Frames are taken as "the fundamental representations of knowledge" and "dynamic relational structures" which are highly context-dependent (Barsalou 1992: 21). A frame is viewed to have a fundamental meaning attribute at both the semantic and the pragmatic levels. *Framing*, accordingly, is referred to as a speaker's applying a frame or a collection of frames to a communicative situation, the process of which involves contextualizing social events in particular patterns of language use (Fillmore 1982a). The practice of framing is, thus, inherently relational and intentional to serve specific purposes in communication, linking human thinking and expectation with language in use.

Frame Semantics distinguishes two different ways of framing. One concerns invoking conceptual knowledge structures to describe situations which may be independent of the immediate discursive context; namely, drawing upon the cognitive prerequisite for understanding meaning (*cognitive frames*) (Fillmore 1982a). The other involves how language speakers conceptualize what is happening in the on-going communication for situated interpretation of roles, purposes, expectations and sequences of action routines

(*interactional frames*) (Fillmore 1982a). The cognitive-communicative duality of framing, as this study shall show, contributes to integrating the individuality and collectivism in terms of knowledge construction through collaborative talk, which would help to ease the tension between rationality and subjectivity in institutional contexts as what has been frequently mentioned in the literature on academic discourse (e.g. Margolis 1992; Tracy 1997).

Outside the field of semantics research, the concept of *frame* in the 1970s was introduced into the field of ethnography and ecological studies for an integrated approach to contextual discourse analysis of social experience and interaction (Tannen 1993). Goffman (1974) defined a *frame* as consisting of principles of organizations which govern social events and human subjective involvement. Framing, accordingly, is understood as ways of communicating in any interactive situations (Goffman 1974). Tannen and Wallat (1993) coined these kinds of frames as *interactive frames* of interpretation to refer to the speaker's sense of activity type and situated meaning-making in communicative contexts (p. 59). Analysis of frame and framing, thus, is an approach to knowledge construction in terms of meaning construction, transmission and interpretation at the interface between human cognition and social interaction (Telles-Ribeiro and Hoyle 2009).

A frame is essentially a larger unit of meaning than what incorporates turns and sequences in interaction and has a potential to include more situational resources for interpretation of the target discourse beyond the immediate context. This has important implications for the research on human mental activities and application of contextual resources in social interaction. The sociological concept of *framing* developed by Goffman (1974) relates "specific linguistic options to the social activity for which language is being engaged" (Drew and Heritage 1992: 9), thus refers to how participants interpret their social practice from

moment to moment, suggesting a conceptual framework to approach context as changing frames of interactional activities.

1.3 Research aims and research questions

As has been outlined, the principal objective of this study is to examine how the NUCASE participants construct knowledge in the setting of university small group talk. More specifically, one aim of the study is to explicate how specific lexical concepts provide access sites to particular cognitive frames so as to build models of disciplinary discourses. This concerns how linguistic forms used by the participants evoke particular frames to describe communicative events based on schematized categories of experience. By addressing this research aim, the analysis will identify basic knowledge domains in each discipline and examine how the participants' relative access to the knowledge domains is established based on their epistemic statuses.

The next level of analysis will address another research aim: to examine how the participants invoke specific aspects of particular cognitive frames to represent meaning in a situated way. This is about how coherence is assigned to the talk by making certain frame elements relevant to the communicative context. The analysis at this level will identify what specific aspects of particular frames are activated through the participants' use of linguistic forms and how semantic potentials of the lexical concepts associated with the linguistic forms contribute to the activation of the knowledge structures. In addition to this comes another research aim: to determine if, and how, the participants' management of sequences at talk would have impacts upon ways of framing-in-interaction for knowledge construction. The analysis at this level will examine how the participants conceptualize what is going on in their group interaction relying on expectations of the conventionalized institutional routines of actions.

Based on the research aims outlined above, three research questions will be answered in this study:

1. What frames are evoked in the key epistemic domains to semantically characterize the disciplinary discourses?
2. How do the participants invoke interpretive frames for situated meaning-making?
3. How do the participants manage framing-in-interaction for knowledge construction in their talk?

1.4 Thesis structure

This study conducts an empirical investigation of the participants' framing behaviours to collaboratively construct knowledge through the talk of disciplinary lexical concepts in small groups. Characteristics of discursive framing are examined at both the epistemological level and the interactional level. There are seven chapters in this thesis. Chapter 1 provides a brief introduction to the background and relevance of the study, articulates the research aims, and presents the research questions. Chapter 2 critically reviews relevant literature on knowledge construction. Chapter 3 introduces the methodological synergy adopted in the study by discussing relevant theories and techniques of Corpus Linguistics and Interactional Linguistics approaches. Chapter 4 focuses on the research design and procedures for data analysis. Chapter 5 reports the research findings based on a comprehensive data analysis. Chapter 6 further interprets and discusses the research findings, addresses the limitations of the research design, and suggests directions for further research. Chapter 7 revisits the research aims, summarizes the main research findings, and outlines the implications and contributions of these findings.

Chapter 2. Literature Review

2.1 Introduction

This chapter reviews relevant theories and previous studies on knowledge construction. Section 2.2 reviewed knowledge construction through human collaboration within three distinct but inter-related dimensions: The linguistic dimension attends to the relationship between language use, knowledge representation, and meaning negotiation. The discourse dimension is about how discourse structures are associated with knowledge forms. The social dimension concerns knowledge construction in disciplinary cultures of academic practice. Section 2.3 focused on the cognitive and communicative aspects of knowledge construction. Frames as knowledge structures for meaning representation and connection are defined and characterized drawing inspirations from the theory of Frame Semantics and research on social interaction. Section 2.4 provided an intensive and critical review of previous research on knowledge construction in the setting of university small group talk.

2.2 Three dimensions of knowledge construction

The status of knowledge has long been faced with a conceptual dilemma between scientific knowledge and socially shared (or ideological) knowledge. The classical epistemology holds the view that knowledge is absolute truth based on the independently removed doubts through reasoning in one's mind. Marx and Engels (1845) drew a distinction between scientific knowledge and ideological knowledge: the former is generated from human life experience in real circumstances, while the latter refers to concepts abstracted from people's life experience that reflect the interests of a particular social class. In terms of acquisition, scientific knowledge involves a method of inquiry which must be based on gathering observable and measurable evidence and subject to specific principles of reasoning and experimentation (Newton 2009). The representation of "the outside world" illustrated in the biological domain

of scientific knowledge emphasizes its nature of being "true" and the general accessibility of the test for the truth. On the contrary, Peirce (1934) argued that there is no such thing as a test for true knowledge and that it is the social community rather than individuals that decides what public knowledge is and whether it can be accepted by the community. The proposition indicates that how people acquire knowledge depends on how they perceive collectively what crucial conceptions mean across social contexts:

> (I)n order to ascertain the meaning of an intellectual conception one should consider what practical consequences might conceivably result by necessity from the truth of that conception; and the sum of these consequences will constitute the entire meaning of the conception.
>
> Peirce (1934: 9)

Peirce's proposition is corresponded by a more recent premise of social interaction that a common ground is needed to make communication possible and meaningful (Clark and Brennan 1991). Benefits of collaboration have been extensively addressed in the fields of anthropology, linguistics, and organizational sciences (see Ellis et al. 1993; Azmitia 1988). Knowledge, along with common memories, experiences, and mental schemas, is shared when people are working together (Hardin and Higgins 1996). The construction of knowledge as an aspect of the shared representation achieved through group work is one mechanism for explaining how collaborative learning happens (Jeong and Chi 1997).

A shared knowledge structure can support collaborative problem-solving among a group of learners by building up a shared conceptual space involving the negotiation of meaning through an "external mediational framework of shared language, situation, and activity" (Roschelle and Teasley 2012: 70). It helps to identify, describe and analyse a problem by

relating previously shared knowledge to available actions for problem-solving in the local situation. Explanations of the relationship make knowledge construction happen and persist into later challenges (Roschelle and Teasley 2012).

The notion of shared knowledge, or *public knowledge* called by Cheung (2009: 40) highlights the accessibility and acceptability of knowledge within a certain social community. Although people in a community may differ from one another in terms of their reasoning preferences, a common ground has to be achieved based on shared experience so that the knowledge can be discussed. This perspective challenges the classical epistemology (see Wolfsdorf 2011[1]), blurring the boundary between "truth" and "falsity", suggesting that knowledge is acquired through discourse-internal negotiation rather than being external to context.

However, how knowledge is efficiently shared and constructed in collaborative settings is far from being "homogeneous and predictable" (Roschelle and Teasley 2012: 94). One cannot expect mutual understanding to be achieved by simply putting people together into a group. As one shall see in the following sections, talk-in-interaction is the most important resource for meaning negotiation to collaboratively construct knowledge; discourse structures have significant influences on how knowledge is to be constructed procedurally; and a broad range of social factors dynamically shape patterns of knowledge construction.

2.2.1 *Linguistic dimension: knowledge, language, and meaning-making*

[1] According to Wolfsdorf (2011:57), there is a close relationship between Plato's epistemology and his metaphysics and ethics, which is reflected from his thought that knowledge requires a rational explanation. From this perspective, the status of being "knowledge" is based on the judgement of an individual mind once doubts are believed to be removed through the process of "reasoning". In such sense, knowledge can only be knowledge only when it can be explained rationally.

The linguistic dimension of knowledge construction concerns the relationship between language, thinking, and meaning-making, which is about how people make sense of the world through the power of language. Knowledge construction, in such sense, can be described as a process of recognizing and interpreting one's socio-cultural environment (Guzmán 2004) mediated by one's language use. Language constitutes "systems for signifying content" (Teubert 2007: 2), through which one can "express, exchange or share a wide variety of knowledge to the world" (Salzmann 1993:153). The important role of language in knowledge construction has been extensively addressed in the literature since "sharing of knowledge is greatly facilitated by language" (Spuzic and Nouwens 2004: 647). Language use and development play a major role in characterizing human culture objectively, both at the local level of interaction and within wider social, political and historical domains (Lantolf and Thorne 2000).

Knowledge is based on "perceptions and thoughts" (Spuzic and Nouwens 2004: 657) which are the most fundamental products of human cognitive activities. One of the most challenging aspects of knowledge construction, however, is the identification of the translation of perceptions and thoughts through language use. The role of language in thinking has been illustrated by Burr (1995):

> (R)ather than viewing language and thought as two separate phenomena which can affect each other, it is suggested that they are inseparable and that language provides the basis for all our thought. It provides us with a system of categories for dividing up our experience and giving it meaning, so that our very selves become the product of language. Language produces and constructs our experience of ourselves and each other.
>
> (p. 30)

Similarly, Gergen (1985) argued that it is through language use that knowledge can be transferred from one individual to another in their interactional practices. This provides the rationale to study knowledge construction by exploring patterns of human language use (Resnick 1987).

Meaning-making is placed at the centre of knowledge construction, especially when this "making-sense-of-the-talk" process is realized through collaboration. Stahl (2006) argued that "the co-construction of shared knowledge in discourse involves the negotiation of tacit meanings" (p. 307). Meaning construction in human everyday activities requires "on-line creativity" and is "negotiated by participants in communication" (Fauconnier 1997a: 8). The negotiation of meaning involving a process of "attunement to the attunement of the other" (Rommetveit 1992: 10) is referred to as:

> (T)he modification and restructuring of interaction that occurs when learners and interlocutors perceive or experience difficulties in message comprehensibility
>
> (Pica 1994: 494)

> communication in which participants' attention is focused on resolving a communication problem as opposed to communication in which there is a free flowing exchange of information
>
> (Gass 1997: 107)

Pica's (1994) statement emphasizes the shaping power of meaning negotiation over the structure of interaction by increasing the comprehensibility of input (see Gass and Varounis 1994; Long 1996). The process of "modification and restructuring", when projected at the action level, includes elaboration, confirmation and comprehension checks, clarification

requests and recasts, etc. (Al Hosni 2014). Gass' (1997) statement, on the other hand, highlights the communicative goal-orientedness of meaning negotiation, suggesting the importance of an interactive learning environment (Long and Robinson 1998) and learners' active participation (Mackey 1999) to ensure that meaning negotiation results in mutual understanding (Ellis 2003).

Scholars in the field of social constructivism (e.g. Vygotsky 1978; Lave and Wenger 1991) argue that the negotiation of meaning is the foundation for all human communication (Nunan 1988) and knowledge construction in collaborative contexts (Bereiter and Scardamalia 2006). *Social constructivism* is a sociological theory of knowledge, a term first coined by Berger and Luckmann (1996). Holding the view that people work together to construct things, social constructivism focuses on an individual's learning in the group interaction. Knowledge from such a perspective is taken as "a product of collaboration" (Cheung 2009: 33), illustrating that language and thought are inevitably linked to each other in our daily experience (Gergen 1985). Different theoretical perspectives have been adopted to address learners' development at a potential level achieved from knowledge assimilation through meaning negotiation (see Vygotsky 1978); or the change of new learners' participation pattern (see Lave and Wenger 1991) which facilitates knowledge construction to create/maintain social coherence of a certain community (Hartnell-Young 2003).

Meaning negotiation relies on the combination of concepts. A concept is commonly perceived as a "unit of knowledge created by a unique combination of characteristics" (SBVR-OMG 2015: 26). Rey (1983) classified two functions of concepts: *linguistic meaning* and *components of cognitive states*. Note that what Rey (1983) meant by linguistic meaning resembles what is regarded as the meaning retrieved from a linguistic form by virtue based on

the semantic relations between lexis. In addition, what was emphasized by Rey (1983) in the function of components of cognitive states is a stance that concepts are cognitively loaded only, understood in a one-way direction to explain human thought and behavior. What has been overlooked in Rey's (1983) categorization is the fact that human linguistic behaviors have the potential of shaping the way concepts are interpreted and developed dynamically in working memory (Barsalou 1987, 1989). In such sense, why and how language users choose particular words in their utterances may be more about how they perceive the selection pragmatically rather than what the words "mean" at the semantic level (see Sweetser and Fauconnier 1996).

In the conceptual system, "a concept takes a form of a *conceptual structure*, knowledge representation assembled for purposes of meaning construction" (Hamawand 2016: 83). When encoded in language, a conceptual structure takes a form of a *semantic structure*, of which meanings associated with words are referred to as *lexical concepts* (Evans 2006; 2009). Some scholars argue that meanings associated with words are protean in nature rather than behaving in a straightforward compositional manner. As suggested by Evans (2006), "the semantic values associated with words are flexible, open-ended and highly dependent on the context" (p. 491; see also Clarks 1983; Sweetser 1999; Coulson 2001; Croft 2000; Evans and Green 2006). Words, therefore, are perceived to have *meaning potentials* instead of meanings:

> (Meaning potentials are) all the information that the word has been used to convey either by a single individual or on the social level, by the language community. A consequence of this approach is that no attempt is made to distinguish between lexical and encyclopedic information in terms of the kind of information that is contained in the meaning potential. Meaning potentials contain both kinds of information— information deriving from the use of language and information deriving from other experience with the world.
>
> (Allwood 2003: 43)

A word, from such a perspective, is viewed as associated with stable linguistic (e.g. morphological, semantic, syntactic, etc.) information but not with the conceptual content by itself; rather, it provides access to conceptual knowledge which is derived from experience (Evans 2009). Word meanings, as represented by lexical concepts, are therefore not "rigid, discrete sense-units, qua mental entities" but projecting relatively unstable knowledge structures which cannot be "straightforwardly identified" (Evans 2006: 493). Words are viewed to have *semantic potentials* to project "situated meaning" through pragmatics to serve specific communicative intentions (Evans 2006: 493). Lexical concepts "represent the semantic pole of linguistic units" in a way that "the specific semantic contribution" is associated with "particular forms" (Evans 2006: 499; see also Langacker 1987). Linguistic devices for meaning construction through lexical concepts are, therefore, believed to evoke knowledge structures for selective interpretation of the ongoing discourse. The properties of lexical concepts as linguistically encoded mental representations are summarized in Table 1.

Property	Details
lexical concepts are form-specific	i.e., lexical concepts are conventionally associated with a specific form. Thus, lexical concepts are necessarily language specific.
lexical concepts are associated with different sorts of forms	i.e., lexical concepts constitute the semantic pole of symbolic assemblies of form-meaning (linguistic units) of all kinds
forms are not lexical concept-specific	i.e., are associated with a 'semantic network' of related lexical concepts, and thus exhibit polysemy
lexical concepts have a lexical profile	i.e., a unique 'biometric' identifier associated with formal and collocational tendencies
lexical concepts can be combined	i.e., lexical concepts can be combined in various predictable ways, due to valence relations (relational vs. non-relational lexical concepts and larger lexical conceptual units), in service of activating semantic potential and thus meaning-construction
lexical concepts possess a semantic value:	i.e., in context, lexical concepts provide an informational characterisation.
i) semantic potential	i.e., access to a cognitive model profile (encyclopaedic knowledge).
ii) encapsulation	i.e., provides access to the cognitive model profile at a particular access site (in the cognitive model profile).
iii) relational vs. non-relational	i.e., relates to specific information concerning valence, and thus combinatorial abilities and properties.
iv) temporal structure	i.e., encodes how the temporal structure of the relation is being accessed, i.e., whether it evolves in time, as encoded by verbs, or whether it is provided as a 'gestalt', as encoded by, for instance, prepositions.
v) referentiality	i.e., lexical concepts refer to or index entities of different sorts. Some lexical concepts refer to entities which are conceived as objectively existing. Other lexical concepts refer to entities understood with respect to some deictic centre, such as the speaker's physical location. Others refer to entities in the linguistic context itself.

Table 1 Summary of properties of lexical concepts (Evans 2006: 511)

2.2.2 *Discourse dimension: knowledge forms and discourse structures*

A great emphasis has been put on understanding knowledge construction through analysing discourse (Cole 1996; Engestrom 1999) since discourse is an essential practice that creates a certain *community* (Wenger 1998) when learners engage in learning activities (Hmelo-Silver 2003). Everyday learning practices have been studied using a broad range of techniques, highlighting the importance of the dynamic relationship between discourse and knowledge (e.g. Cazden 1986; Cobb and Yackel 1996; Koschmann et al. 2000). According to Koteyko (2006), discourse provides "several ways to see how meaning is created in language use" by implying "a complex relationship between the linguistic and the social, and the different

approaches construe this relationship on different terms" (p. 132). Structural configurations of discourse are closely associated with meaning production for knowledge construction.

Discursive skills are, therefore, regarded as one of the key competences (Heller and Morek 2015) for knowledge construction because they enable learners to participate in learning activities across different subjects and disciplines (Uccelli et al. 2014). The view that knowledge is the product of a constructive discourse process rather than the by-product of communicative activities is shared by Bereiter and Scardamalia (1989). They initiated a model of *knowledge building* which was later developed into the term *knowledge creation*. Referred to as "a set of social practices that advance the state of knowledge within a community over time" (Paavola et al. 2004, cited by van Aalst 2009: 260), the term emphasizes the status of discourse in determining the limits of knowledge, setting goals, investigating problems, promoting the impact of new ideas and evaluating whether the state of knowledge is advancing. Similarly, van Aalst (2009) distinguished three modes of discourse:

- *Knowledge sharing* discourse refers to the transmission of knowledge between people, involving introducing information through collaborative inquiry in a community with a lack of attention to interpretation, evaluation, and development;
- *Knowledge construction* discourse refers to the processes by which students solve problems and construct understanding of concepts, phenomena, and situations, featured as effortful, situated, reflective at the interface of being individual and social (Palincsar 1998);
- *Knowledge creation* discourse: involves the design and improvement of intellectual artifacts such as theories, explanations, and proofs (Bereiter 2002), emphasizing explanations, casual mechanisms, and the coordination of claims and evidence.

(p. 260-264)

Knowledge construction is often associated with *deep learning* in a community (Biggs 1987), involving "qualitative changes in the complexity of students' thinking about and conceptualization of context-specific subject matter" (Moore 2002: 27). Knowledge construction contributes to restructuring knowledge in a substantial manner (van Aalst 2009) at different levels (usually from moderate to high) of *engagement* - a term coined by Dole and Sinata (1998), referred to as the effort a learner makes to process information through learning practices. Cognitive activities are involved in knowledge construction processes which would imply deeper insights into both the concepts themselves and the networks of concepts (van Aslst 2009). New concepts can be invented and enhanced during knowledge construction processes as well, along with the practice of interpreting, evaluating, critiquing, testing and sharing these concepts within and outside a community.

As argued by some scholars, knowledge construction is placed at a higher level than knowledge sharing (e.g. Hmelo-Silver et al. 2007) through processes directed at the acquisition of the reliable knowledge of a field (see Kolodner et al. 2003; Krajcik et al. 2008). The intensive focus on sharing knowledge by increasing informational input may explain why many educational approaches do not perform well in practice. Some ill-structured learning tasks make learners misunderstand knowledge as isolated, segmented, static "pieces of things" they could quickly pick up to solve the problem provided to them. Such a "to-know-for-the-sake-of-knowing" orientation to knowledge has to be replaced by a new model which goes beyond the acquisition level into broader conceptual and practical domains. Such a new model shares the features of Bereiter and Scardamalia's (1996) *knowledge creation* model, following the principle that discourse plays an important role in enhancing knowledge innovation by "maintaining social relations, setting goals, deepening inquiry and lending support to ideas that are already understood by some in the community" (van Aalst 2009:

263). In such sense, knowledge construction and knowledge creation are overlapping, consecutive or even interwoven processes in a community's learning practices. A re-conceptualization of knowledge construction is needed and the discursive practices through which people deal with knowledge should be viewed in a dynamic circle rather than a liner order, in a systematic rather than an isolated manner.

Specifically, to explore knowledge construction in the educational context, I found Bernstein's work (1971; 1999) particularly relevant. In his *Classification and Framing* paper (1971), Bernstein introduced the term *educational knowledge code*, whose forms depend on "social principles which regulate the classification and framing[2] of knowledge made public in educational institutions" (p. 47-48). He explained the two terms *classification* and *framing* as follows:

- *Classification* refers to the degree of boundary maintenance between contents… Classification is strong when contents are well insulated from each other. Classification is weak when the boundaries between contents are blurred.
- *Framing* refers to the form of context in which knowledge is transmitted and received, the specific options of power distribution and discourse control which shape the specific pedagogical relationship. Strong framing entails reduced options, while weak framing entails a broader range of options.

(p. 47-48)

Bernstein's classification does not refer to how disciplinary contents are classified, but the degree of insulation between them in pedagogical interaction. As illustrated by Bernstein

[2] The term *framing* coined by Bernstein (1971) in his educational knowledge code is defined from a pedagogical perspective. The focus is on knowledge transmission and reception in formal educational institutions, which is different from how frame and framing are conceptualized in Frame Semantics and Interactional Linguistics.

(1971), very strong insulation can be identified between the 'pure' and the 'applied' knowledge (see Becher 1987a, 1987b) in England at levels below Higher Education. It should be noted that there is difference between educational levels concerning how 'pure' and 'applied' disciplines are to be perceived. At levels below Higher Education, a pure variety of subjects are drawn from a common universe of knowledge (e.g. chemistry, physics, mathematics) while an impure variety of subjects are drawn from different universes of knowledge (e.g. religion, physics, economics) (Bernstein 1971, original examples). The optional combinations as such are, to a large extent, in the form of the *collection knowledge codes* (Bernstein 1971: 51), where disciplinary boundaries are quite clear and relatively closed. On the other hand, Higher Education manifests features of the *integrated knowledge codes* (Bernstein 1971: 53) where disciplines are developed into a more intricate structure bound by some relational ideas, thus with blurred boundaries.

It should be noted that *framing* in Bernstein's (1971) educational knowledge code is different from how frame and framing are conceptualized in semantics research (see Section 2.3.1) and social interaction studies (see Section 2.3.3). It does not refer to *what* knowledge is for itself, but *how* knowledge is constructed; or in Bernstein's (1971) words "ways of knowing" (p. 57). *Frame* in Bernstein's term represents patterns of power distribution and control of institutional discourse in contextualized pedagogical relationships. Bernstein (1971) identified a weaker frame in the teacher-pupil relationship in England compared with that in Europe, while understanding of framing of knowledge in the Higher Educational context is to a large extent limited, if not missing in his work.

In his 1999 paper, Bernstein distinguished two modalities of discourses *vertical discourse* and *horizontal discourse* with "forms of knowledge" (p. 158) as the criteria. He argued that

different forms of knowledge will be realized and constructed within the two types of discourses:

- *Vertical discourse*: A vertical discourse takes the form of a coherent, explicit, and systematically principled structure, hierarchically organised, as in the sciences, or it takes the form of a series of specialised languages with specialised modes of interrogation and specialised criteria for the production and circulation of texts, as in the social sciences and humanities.
- *Horizontal discourse*: A horizontal discourse entails a set of strategies which are local, segmentally organised, context specific and dependent, for maximising encounters with persons and habitats.

Bernstein (1999: 159)

Defining features of the two types of discourse are shown in Table 2:

TABLE II

	Vertical discourse	Horizontal discourse
Practice	Official/institutional	Local
Distributive principle	Recontextualisation	Segmentation
Social relation	Individual	Communalised
Acquisition	Graded performance	Competence

Table 2 Features of vertical and horizontal discourse (Bernstein 1999: 162)

Based on the distinction between vertical and horizontal discourses, Bernstein further argued that these two types of discourses entail different modes of learning and context management patterns due to different structures of knowledge acquisition. His main propositions are summarized as follows:

Both vertical and horizontal discourses are likely to operate with distributive rules that set up positions of defense and challenge. In vertical discourse, there are strong distributive rules regulating access, transmission and evaluation. Circulation of knowledge is basically realized through explicit re-contextualization and evaluation which are motivated by strong distributive procedures. In horizontal discourse, there is inter-dependence between the social relationships and the forms of discourse. Circulation of knowledge is basically realized through the segmental organization based on its contextual mode of orientation to social solidarity. The mode of knowledge acquisition is created by the form taken by the pedagogy, the interventions of which shape the form of knowledge to be acquired. In vertical discourse, knowledge is mainly acquired through an integration of meanings based on specialized symbolic structures of explicit knowledge. The procedures for the acquisition are linked hierarchically. Development is seen as the development of theory. In horizontal discourse, knowledge is mainly acquired through the relevance of one contextual segment to another. A general competence rather than a graded performance is acquired. The competences are segmentally related, context-specific and dependent, highly goal-oriented to the on-going practices. Development is seen as the introduction of a new language to offer the possibility of fresh perspectives, new questions, new connections and a new set of speakers.

Bernstein (1999: 160-163)

While Bernstein's work on knowledge and discourse structures is illuminating, it is limited in the following aspects:

1) There has been little first-hand information drawn from empirical studies to test the validity of the framework.
2) The investigations have been mainly conducted within educational practices below Higher Education, with little research specifically focusing on university settings.
3) Generic specifics within the educational register have been to a large extent ignored (e.g. formats of classroom organization, types of participants, language varieties, etc.).
4) Products rather than processes of knowledge construction have been emphasized, with the interactional mechanism in collaborative learning contexts reduced to trivial.

In sum, Bernstein's (1971; 1999) work has provided a useful way of approaching the relationship between knowledge construction and discourse structure in educational practice. This inspires the design of the current study to address the research focus (collaborative knowledge construction) within the discourse dimension. On the other hand, due to the limitations mentioned above, the current study has been conducted in a setting (university small group talk) with the research focus shifted to the processes rather than products of knowledge construction and a broader scope of contextual factors taken into account.

2.2.3 *Social dimension: knowledge construction and academic disciplines*

Since the 1980s, sociological perspectives on epistemology have made explicit the great importance of the social context in knowledge construction (Kittleson and Southerland 2004). For example, Lave (1988) introduced the concept of *situated cognition* to understand knowledge as being "stretched across mind, body, activity and setting" (p. 18). Hutchins (1995) called for a fuller view of knowledge and cognition beyond the laboratory research setting. Meanwhile, sociocultural theories posit knowing as the ability to participate in various cultural practices (Lave and Wenger 1991) and draw upon the relevance of language and thought by describing the former as "a social mode of thinking" (Mercer 1995: 4) to shape knowledge construction. Knowledge has been viewed to exist as "a social entity" rather than "an individual possession"[3], whose essence lies in the fact that "it is shared" (Mercer 1995: 66).

Knowledge is acquired through social interaction where higher mental functions (other than what was biologically endowed) are developed with significant interventions of other

[3] From such a perspective, the "individual" social relation entailed in vertical discourse suggested by Bernstein (1999) may need to be re-examined.

people/tools in one's life. The higher-level cultural tools include language, literacy, numeracy, categorization, rationality and logic (Lantolf and Thorne 2000:198). The use and meaning of such tools (along with the physical ones) are first externalized at the interpersonal level through interaction and then internalized at the intra-level (Vygotsky 1978). Social interaction helps the participant to derive meaning from the world and contributes to shaping his/her habits of mind for meaning production. Certain concepts are "re-externalized" (Arvaja 2007: 133) through collaborative knowledge construction when new concepts and understandings are built (Stahl 2004). The specific knowledge constructed this way within a certain community of social practice, in turn, represents the shared value, collective identity and re-mediated knowledge orientation among its members.

At the action level, participants are found to index their *relative epistemic positions* (Heritage 2013: 556) through sequences of social interaction (see also Hayano 2011; Heritage 2011; Heritage and Raymond 2005, 2012; Raymond 2003; Raymond and Heritage 2006; Stivers 2005). A more fully sociological conception of *knowledge domains* (Berger and Luckmann 1967; Lukes 1974; Abbott 1988) had been introduced since Kamio (1997) developed the concept of *territories of information* to explain the relativistic conception of access to knowledge. As he suggested, information known by each speaker in interaction can be arrayed on a continuum from what is highly distant from the speaker to what is extremely close and can be taken as possessed by the speaker. The conceptualization not only involves who knows what and how, but also who has the right, responsibility and obligation to know it and articulate it (Kamio 1997; Maynard 2003; Terasaki 2004; Raymond and Heritage 2006; Stivers et al. 2011). This concerns moment-to-moment expression and negotiation of social relationships, which indicates that knowledge construction in social interaction is a highly fluid and dynamic process.

In educational settings (especially for the ones in Higher Education), knowledge construction is mainly realized through the use of academic language. Academic language refers to a unique type of linguistic register assumed to be functional for the purpose of knowledge production and transmission through teaching and learning activities (Heller and Morek 2015). Academic language has been taken as a medium of knowledge construction and a tool for thinking with socio-symbolic functions to shape individual identities and patterns of community practice (Wenger 1998). At the linguistic level, the register-based approach (e.g. Halliday 1978) to academic language, cognition and content learning emphasizes the functional value of certain linguistic forms which is aimed to present information "in highly structured ways" to enable the language user "to take an assertive, expert stance toward the information presented" (Schleppegrell 2001: 451).

According to Halliday (1993), linguistic features typical of academic language are contrastive to mundane language use by representing a very specialized way of thinking about the academic world rather than merely talking about the same thing in the general world using another language system. Constructing knowledge through the use of academic language is of particular interest as being socialized (Hyland 1999) within disciplinary discourses (Becher 1987a). An epistemic perspective has been adopted to view academic disciplines as

> defined and differentiated by distinct and often radically different conceptions of what knowledge is, where it is located, how it is to be constructed, and what the process of knowledge construction is for
>
> (Groom 2007: 1-2)

Academic disciplines, therefore, can be characterized as distinct territories of information, or *epistemic domains*, a term coined by Stivers and Rossano (2010). Different from those involved in mundane conversations, epistemic domains differentiated by academic disciplines see more of "the stability of what occurs as a consequence of the social context" (Tannen and Wallat 1987: 206) with meanings to a larger extent shaped by participants' shared understanding of institutional history and practice. This means that the rights and obligations to claim the access to particular knowledge may be "for the most part presupposed or agreed upon" (Heritage 2012: 6). Some scholars, thus, assume that different disciplines feature different knowledge structures and skills (Nishina 2010). For example, "descriptive skills" (e.g. to describe apparatus, procedures or findings) are prioritized in computer science, engineering, while "argumentative skills" (e.g. for a particular theoretical position) are essential in psychology, sociology and business studies (Bridgeman and Carlson 1984). Typically, disciplines of science and technology often require activity-based skills while disciplines of social sciences and humanities value inquiry techniques drawing upon multiple sources and synthesis methods (Casanave and Hubbard 1992).

These assumptions have important implications for understanding the epistemological aspect of disciplinary discourse (Berkenkotter and Huckin 1995); namely, the shaping power of knowledge orientation in education which contributes to the establishment of "disciplinary communication systems" (Stichweh 2003: 1). In his exploration of the ways in which academic citation practices contribute to the construction of disciplinary knowledge, Hyland (1999) found that *hard knowledge* (e.g. biology, electronic engineering, mechanical engineering, physics, etc.) features a "relatively steady cumulative growth" where "new findings are generated by a linear development from an existing state of knowledge" (p. 352), with the research "proceeding along well-defined paths" (p.352). On the other hand, such

predictability and routineness are relatively rare in social sciences and humanities where *soft knowledge* (e.g. philosophy, business, sociology, applied linguistics, etc.) is constructed in more "reiterative and recursive routes" (p. 353). Nevertheless, it is rather common that participants frequently "retrace others' steps and revisit previously explored features of a broad landscape" (Becher 1989: 353) to construct new knowledge from prior knowledge, no matter they follow a "hard" or "soft" pattern. Hyland (1999) also suggested that language use to construct soft knowledge is "open to greater interpretation", with findings "more frequently borrowed from neighbouring areas" and "no clear-cut criteria for establishing for refuting claims" (p. 353).

Among the epistemologists who have been striving to categorize academic disciplines based on types of knowledge and skills, Becher (1987a, 1987b, 1989, 1994) and his colleagues (Becher and Trowler 2001; Neumann et al. 2002) have done the best-known work in the field. Becher developed a theory of *knowledge domains* based on the nature of knowledge. Four knowledge domains are termed in his categorization to form a two-dimensional space where individual disciplines can be posited in an axis (Figure 1, original examples):

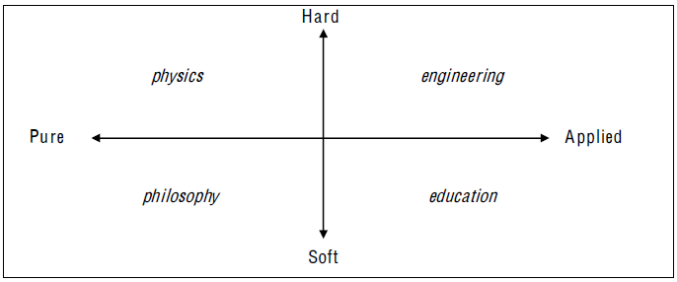


Figure 1 Becher's knowledge domains (adapted from Groom 2007: 18)

What is worth noticing in Becher's work is his explication of the relationship between the nature of disciplinary knowledge and the discursive patterns regulated by different linguistic

practices. According to him, pure science (e.g. physics) is based on a process of "accretion of knowledge" and hence may be taken as "atomistic"; while soft disciplines (e.g. history, sociology) feature inquiries based on "criticism, reiteration and reinterpretation", the knowledge of which is thus "organic, holist, complex and comparatively unpredictable" (Becher 1987a: 269-207). As a result, two distinct discursive patterns are identified: a bottom-up process for the former by assembling structures and components piece by piece and a top-down process for the latter from a relatively holistic perspective (Becher 1987a).

Disciplinary grouping	Nature of knowledge	Nature of disciplinary culture
Pure sciences (e.g. physics): 'hard-pure'	Cumulative; atomistic (crystalline/tree-like); concerned with universals, quantities, simplification; resulting in discovery/explanation.	Competitive, gregarious, politicall well-organized; high publication rate; task-oriented.
Humanities (e.g. history) and pure social sciences (e.g. anthropology); 'soft-pure'	Reiterative; holistic (organic/river-like): concerned with particulars, qualities, complication; resulting in understanding/interpretation.	Individualistic, pluralistic; loosely structured; low publication rate; person-oriented.
Technologies (e.g. mechanical engineering): 'hard-applied'	Purposive; pragmatic (know-how via hard knowledge); concerned with mastery of physical environment; resulting in products/techniques.	Entrepreneurial, cosmopolitan; dominated by professional values; patents substitutable for publications; role-oriented.
Applied social sciences (e.g. education): 'soft-applied'	Functional; utilitarian (know-how via soft knowledge); concerned with enhancement of [semi-] professional practice; resulting in protocols, procedures.	Outward-looking; uncertain in status; dominated by intellectual fashions; publication rates reduced by consultancies; power-oriented.

Table 3 Knowledge and culture by disciplinary groupings (Becher 1987a)

Table 3 shows Becher's model of variations across disciplinary discourses based on the nature of knowledge and disciplinary cultures. During the last decades or so, this categorization has provided essential theoretical underpinnings for the research in Applied Linguistics concerning disciplinary differences in academic writing and publication (see Hyland 2000; North 2005; Charles 2006). The popularity of Becher's work lies in its explication of the relation between disciplinary epistemologies and linguistic practices of academic writers and speakers (Groom 2007). Such a hypothesis has a conceptual significance that institutionalized

language behaviours represent knowledge structures as a cognitive entity, or as Becher (1987a) himself put it:

> If (my) general thesis … is tenable, one would expect differences in fields of knowledge to be reflected in differences in linguistic form: and by the same token, differences in linguistic form to signify differences in fields of knowledge.
>
> (p. 261)

Becher's categorization of academic disciplines, as far as I see, benefits the research of knowledge construction through academic language use in the following aspects:

- It creates two continuums (horizontally and vertically) along which disciplinary features reflected in participants' language behaviours can be investigated, compared and evaluated;
- It identifies the most basic parameters (hard vs. soft; pure vs. applied) to be controlled for legitimate cross-disciplinary comparisons and corpus homogeneity enhancement;
- It relates the nature of disciplinary knowledge/skill to linguistic practices and discourse structures it is embedded in thus provides a pass to explore knowledge/skill-oriented interactional strategies for academic purposes.

However, Becher's work has been exclusively based on written language (e.g. journal articles) and on the relatively "formal" scholar communication in academic communities. It focuses on participants' individual (writing) performance rather than their competence in oral interaction, thus risks incompleteness and oversimplification. There has been considerable focus on the social construction of knowledge at the macro level in Becher's categorization, yet insufficient attention has been paid to "knowledge in communication" (Stivers et al. 2011: 6; see also Levinson 2012) in the framework. What is elusive in Becher's work is a systematic

investigation of the mechanism of knowledge construction at the micro level. The categorization at the macro level, though articulate and useful, can only serve as a starting point from where the researcher may go further to highlight the implicit (e.g. academic identity, discourse control) in addition to the explicit (e.g. rules and regulations, program development) institutional impacts on knowledge construction.

A great effort has been made to identify and analyse disciplinary differences concerning participants' professional identities and discursive decisions (e.g. Beijaard et al. 2004; O'Connor 2008; Eliot and Turns 2011, etc.). The existing studies on academic language use for knowledge construction, however, suffer serious shortcomings and limitations. Their focus has been intensively put on systematic and generalizable features of certain registers and linguistic forms with a predominant attention to written texts as the empirical basis for examination, description, and characterization (Heller and Morek 2015). A shift of view towards a contextualization perspective with more emphasis on spoken academic language use is urgently needed. The rationale of looking at collaborative knowledge construction in academic talk at the university level has been emphasized by Atwood et al. (2010):

> Given that the university is considered a place where reasoning is highly valued and where complex ideas are developed, the quality of talk constituting the knowledge-building process at the university level deserves specific attention.
>
> (p. 363)

The statement above suggests the inextricable connection between academic disciplinary knowledge and the language use in its modes of interaction. The distinctiveness of ideas, research problems, theoretical propositions, paradigms, presentations and evaluations of

knowledge (Clark 2003) are all implied by distinct ways of "speaking of them". For example, 'I don't know' in mundane conversations has been widely accepted as a claim of insufficient knowledge (Beach and Metzger 1997) to mark an unknowing epistemic stance in response to a question (K ärkk änen, 2003). However, Park (2012) found that the lexical bundle also functions as a 'prepositioned epistemic hedge' to invoke the epistemic primacy of the teacher when the student uses it to downgrade his/her own epistemic veracity and to show less commitment to his/her following talk (Weatherall 2011: 317).

Investigation of verbal linguistic features of disciplinary discourse can provide meaningful insights on knowledge construction and its sociocultural implications (Nishina 2010) since words in disciplinary talk are schematically interrelated to both the meaning and the reality of a particular discourse community (Swales 1990; Widdowson 1998). The linguistic practice of such a community to engage its members (Hyland 2002) can to a large extent reveal how knowledge and language mutually shape each other to achieve social coherence.

2.3 Frame and framing for knowledge construction

As discussed in Section 2.2.1, lexical concepts associated with words as linguistically encoded mental representations are not units of meaning by virtue but provide points of access (Langacker 1987; Evans 2009) to large-scale conceptual knowledge structures. Word meanings are viewed as a consequence of these coherent knowledge structures constructed from a range of ready-made schemas which underline "the mutability of human experience" (Evans 2006: 497). Such structures are referred to as *frames*, conceptual entities defined by lexical concepts through linguistic encoding and evoked by given words in discourse. Marvin Minsky first used the term *frame* for knowledge representation in his 1974 article *A framework for representing knowledge*. He was interested in "a theory of human thinking"

(Minsky 1975: 215) and described a *frame* as "a data-structure for representing a stereotyped situation" (p. 3). According to him, attached to each frame are several kinds of information, some of which is about how to use the frame, some about expectations of what happens next and some about what to do if the expectations are not confirmed.

The coinage of the term *frame* was taken as a response to the insufficiencies in the established theories of meaning and knowledge which were "too minute, local, and unstructured to account – practically or phenomenologically—for the effectiveness of common-sense thought" (Minsky 1974: 1). Minsky (1974) proposed the search for a unit of knowledge representation, or in his own words, "the chunks of reasoning, language, memory and perception" which are expected to be "larger and more structured" with connected "factual and procedural contents" to explain "the apparent power and speed of mental activities" (p. 2). The idea was shared by other scholars working on theories of knowledge and intelligence, including Newell and Simon's (1972, *Problem Space*), Schank (1972), Abelson (1973), Schank and Abelson (1975, *script*), Norman (1972), Lakoff (1977, *Idealized Cognitive Models*, also known as ICM), Rumelhart (1980, *schemata*) and also himself (Minsky 1969). The essence of Minsky's theory is articulated as follows:

> When one encounters a new situation (or makes a substantial change in one's view of the present problem) one selects from memory a structure called a Frame. This is a remembered framework to be adapted to fit reality by changing details as necessary.
>
> (Minsky 1974: 1, original emphasis)

The basic structure of a frame as described by Minsky (1974) is a network of *nodes* and *relations*:

The top levels of a frame are fixed and represent things that are always true about the supposed situation. The lower levels have many terminals -- slots that must be filled by specific instances or data. Each terminal can specify conditions its assignments must meet… The frame systems are linked, in turn, by an information retrieval network… These inter-frame structures make possible other ways to represent knowledge about facts, analogies, and other information useful in understanding.

Minsky (1974: 1-2)

According to Minsky (1977), knowledge is stored in "collections of related frames" (p. 355). Although the dynamics and flexibility of knowledge construction has been emphasized, Minsky's (1974) theory of frame is not necessarily unified or coherent concerning how exactly frames as memorized structures are retrieved in different situations. What remains unspecified is the boundary of the flexibility existing in the inter-frame structures which is later believed to be determined by the interpretation of contextual cues. On the other hand, Minsky tended to rely on the power of his own hypotheses and the power expected from the reader for frame application and interpretation, which is to a large extent through the imagination of the proposed cognitive processes (Bednarek 2005). This theory of frame, in its initial form, is open to questions, challenges and at the same time the potential for development since it is incomplete and fragmented, just as Minsky himself admitted:

(T)he schemes proposed herein are incomplete in many respects. First, I often propose representations without specifying the processes that will use them. Sometimes I only describe properties the structures should exhibit. I talk about markers and assignments as though it were obvious how they are attached and linked; it is not.

Minsky (1975: 213)

2.3.1 *Frame and framing in Frame Semantics*

The notion of *frame* has seen further interpretation and development in the field of *Frame Semantics* where the work of Charles, J. Fillmore (1961; 1968; 1969; 1971; 1975; 1978; 1982a; 1982b; 1985a; 1985b; 1986; 1987) and his colleagues (Fillmore and Baker 2001; 2009) is particularly influential. Fillmore initiated Frame Semantics as a theory of meaning in the 1980s to propose "a research program in empirical semantics and a descriptive framework to present the results of such research" (Fillmore 1982a: 373). The theory holds the view that word meaning is characterized in terms of experience-based schematizations of the speaker's world (Petruck 1996). A *frame* was defined by Fillmore (1975) as "any system of linguistic choices… which can get associated with prototypical instances of scenes (p. 124). He then interpreted the term in a more cognitive manner - "a system of categories structured in accordance with some motivating context" (Fillmore 1982a: 381). Frames as constructs derived from "an approach to language as a system of communication" provide a way of "characterizing the structured encyclopedic knowledge which is inextricably connected with linguistic knowledge" (Cienki 2012: 1). Table 4 shows the principles and properties of the Frame Semantics theory.

Theory / Parameter	Frame Semantics
Nature of theory	A theory of meaning; an encyclopaedic theory which takes both linguistic and non-linguistic phenomena into account (Hamawand 2016).
Basic proposal	Lexicon has a frame structure.
Core terminology	Semantic frame: "a data-structure for representing a stereotyped situation" (Minsky 1974: 3); "a system of categories structured in accordance with some motivating context" (Fillmore 1982: 381); "a knowledge structure" based on "recurring human experience" which "relates lexical items linked with a particular scene" (Hamawand 2016: 146).
Structural feature	Embedding both knowledge itself and how to use the knowledge (Rumelhart 1980); dynamic relational structures which is context-dependent (Barsalou 1992); an interpretative device (Fillmore 1985); describing a type of event, relation, or entity and the participants in it (FrameNet Project 1997).
Meaning claim	The meaning of words can only be fully (be best) understood with reference to a structured background of experience, beliefs or practices (Fillmore and Atkins 1982; Kittay and Lehrer 1982; Barsalou 1992; Fauconnier 1992, 1997; Fauconnier and Turner 2006, etc.).
Relation to context	Highly context dependent.
Fundamental	Semantics and pragmatics.
Traditional usage	To characterize word meaning in terms of experience-based schematizations of the speaker's world (Petruck 1996).

Table 4 Principles and properties of Frame Semantics

That a frame is defined as a knowledge structure for mental representation has been widely accepted by linguists. Such knowledge structures are "not innate but acquired through socialization as constructed out of experience", thus are highly "culturally dependent" (Bednarek 2005: 690; also see Tannen 1993; Yule 1996). Hamawand (2016) describes a frame as "a knowledge structure" based on "recurring human experience" which "relates lexical items linked with a particular scene" (p. 146). Barsalou (1992) proposed that "frames provide

the fundamental representation of knowledge in human cognition" (p. 21); and that "frames are dynamic relational structures whose form is flexible and context dependent" (p. 21). A frame, in Barsalou's (1992) view, includes "a co-occurring set of abstract attributes that adopt different values across exemplars" (p. 23). Framing, accordingly, is understood as a speaker's applying a frame or a collection of frames to a communicative situation for specific purposes. The speaker intends the frame(s) by addressing lexical concepts which are "recognized as grounded in" the frames (Fillmore 1982a: 382). The general concept of framing, as argued by Fillmore (1982a), "involves contextualizing or situating events in the broadest sense possible" concerning established "patterns of framing" specifically associated with "given lexical items or grammatical categories" (p. 391).

Fillmore (1982a) emphasized the necessity of distinguishing two different ways cognitive frames get involved in the interpretation of discourse. On the one hand, there are cases in which textually observable lexical and grammatical materials *evoke* the relevant frames in the mind of the interpreter since the former exist as "indices" of the latter; on the other hand, there are cases in which the interpreter "assigns coherence to a text by *invoking* a particular interpretive frame" (p. 385). Invoked frames can come from "general knowledge", knowledge existing "independently of the text at hand" or "the ongoing text itself" (Fillmore 1982a: 386). However, Fillmore did not explain the relationship between frames evoked by linguistic forms and frames invoked for text interpretation when both ways of frame application are involved in a particular context.

Fillmore (1982a) also distinguished two types of framing. One includes structures of motivating categories speakers wish to bring into play when describing situations that might be independent of the actual speech situation. This type of framing concerns the cognitive

aspect of knowledge structures and framing activities, emphasizing the "conceptual prerequisite for understanding the meaning" evoked by *cognitive frames* (Fillmore 1982a: 376). The second type of framing is the framing of the actual communicative situation, through what is called by Fillmore (1982a) as *interactional frames*. This kind of framing has to do with how people conceptualize what is going on between the speaker and the hearer in communicative contexts, concerning an abstract structure of expectations to define the roles, purposes, conventionalized sequences of language-in-action associated with certain frames (Fillmore 1982a).

Cognitive frames (Fillmore and Atkins 1992: 76) are often perceived as interchangeable with *knowledge schemata* which was coined in the field of Artificial Intelligence and later introduced into the field of Discourse Analysis. Cook (2009) defines *knowledge schemata* as "mental representations of typical situations…used in discourse processing to predict the contents of the particular situation which the discourse describes" (p. 69). Such conceptual structures can be perceived as clusters of knowledge containing sets of concepts and associations among the concepts, and are important to interpret intentions, leading to what is called *expectation driven understanding* (Cook 2009: 71) of discourse. Such understanding is based on logical relations between units of meaning, thus contributes to explaining the speaker's lexical choices to arrange information in communication (Cook 2009).

During the past four decades or so, research on frames and framing specifically in the field of linguistics studies has been so drastically developed that the predictability of frame applications in real language use becomes salient. What is worth noticing is the FrameNet Project which has been in operation at the International Computer Science Institute in Berkeley since 1997. The FrameNet Project has been developed as a lexical database of

English that is both human- and machine-readable, based on annotated examples of how words are used in actual texts. The main dataset for the example collecting is the *British National Corpus* (BNC). FrameNet is a lexical resource for English that records the semantic and syntactic valences of each lexeme in terms of Frame Semantics (see Ruppenhofer et al. 2006). Figure 2 and Figure 3 below show a glimpse of the interface of Frame Index and Lexical Unit Index on the FrameNet website.

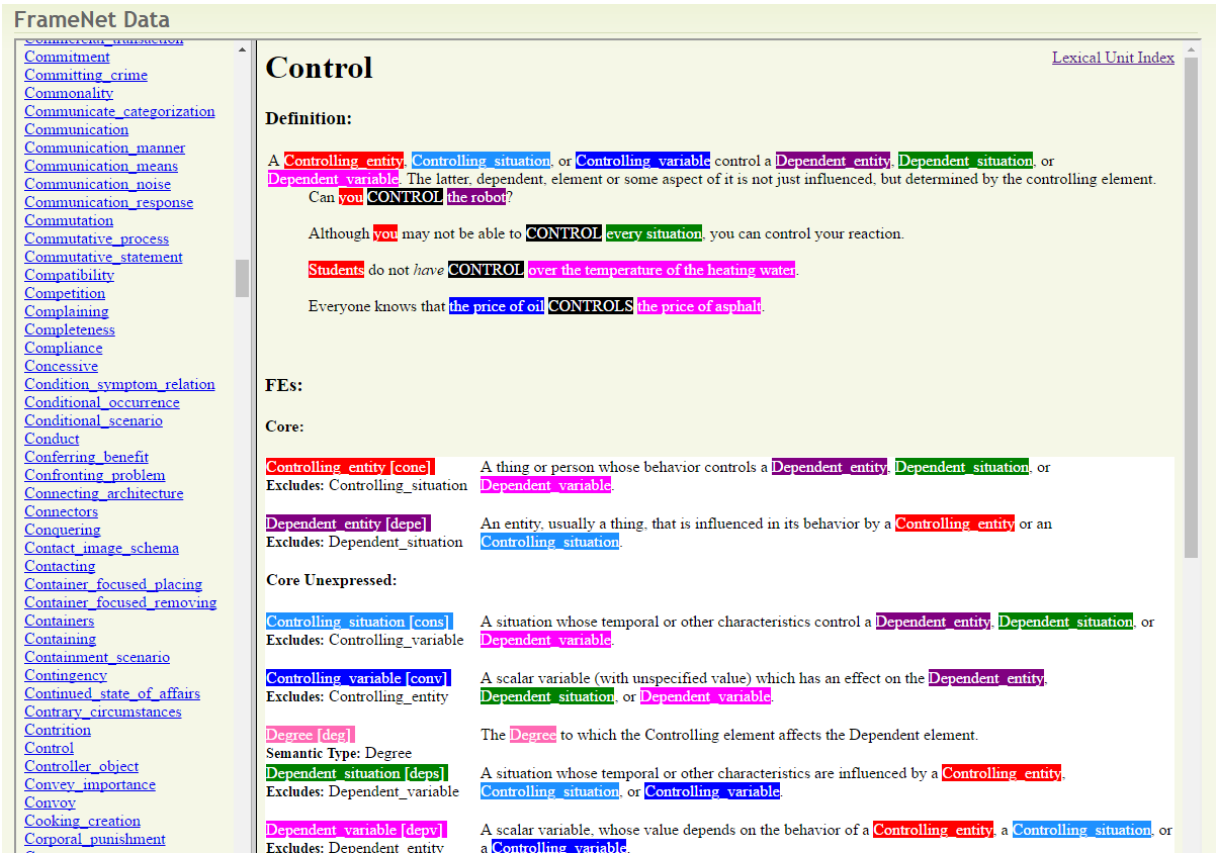


Figure 2 FrameNet Project website_cognitive frame of CONTROL

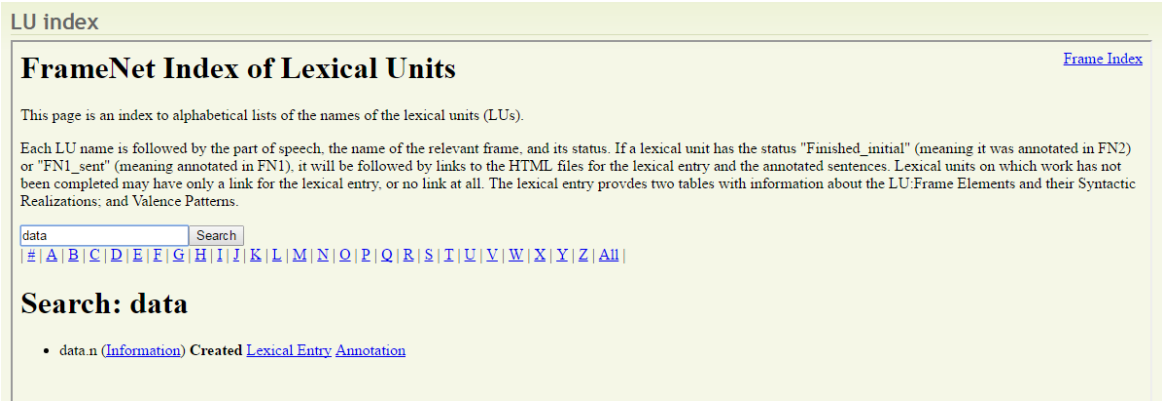


Figure 3 Lexical Unit Index on the FrameNet Project website

FrameNet Project as an external reference is of particular relevance to this study, concerning how the participants' framing behaviors can be methodologically approached based on certain typology of cognitive frames. FrameNet Project has indexes and lists of frames, lexical units and frame lattice on the website, which provides a good starting point to for the researcher to examine valid categorization of frames from cognitive perspectives.

According to Fauconnier and Turner (2006), "frames structure our conceptual and social life… Words are themselves viewed as constructions, and lexical meaning is an intricate web of connected frames" (p. 303). In Frame Semantics, the meaning of a word is understood in relation to its background frame, not in relation with other words (Fillmore 1985). In other words, a word's meaning depends on its conceptual underpinning, knowledge of which is necessary for its appropriate use.

Pragmatics is, therefore, fundamental to Frame Semantics in a sense that people not only depend on the perception of frames but also strive towards applying frames in creative ways to achieve particular communicative goals. Frame Semantics draws no distinction between

semantic and pragmatic knowledge. It is argued in the theory that the meaning of a word subsumes knowledge both of what it is and how it is used and that pragmatic meaning is the real meaning of a lexical item (Hamawand 2016: 149). Similarly, Fauconnier (1997) argued that configurations of conceptual structures in discourse "are built up semantically and pragmatically at the same time" (p. 70). According to him, there is "no configuration corresponding only to the semantic information that would later be patched up by the pragmatics" (Fauconnier 1997: 70). In spite of the divisions between semantic and pragmatic constructions, the two are inseparable in inquires of frames and framing.

As argued by Fillmore (1977), "if a lexical item exits, in other words, it must exist as some part of a frame and must correspond to some part of a schema" (p. 135). Frame Semantics has thus closely linked modern cognitive semantics and more distinct ones with structural perspectives, attending to the pragmatic and contextual encyclopedic knowledge rather than merely focusing on the semantic aspect of word meaning (Nerlich and Clarke 2000). The theory emphasizes that conceptual schemata and situated categorization of knowledge are contextually activated by linguistic expressions in real language use. Interpretation of meaning is therefore not necessarily dependent on the interrelationship between words belonging to particular semantic fields but is a cognitive consequence to a broader conceptual realm where the word is embedded in (Nerlich and Clarke 2000).

The notion of *frame* highlights "the semantic supporting function of domains for concepts" (Clausner and Croft 1999: 1). Concepts are perceived as underlying supporting structural elements of frames. In the conceptual system, "a concept takes a forms of a *conceptual structure*, knowledge representation assembled for purposes of meaning construction" (Hamawand 2016: 83). Frame Semantics views the understanding of word meanings as

realized within conceptual structures. To quote Lowe et al. (1997):

> In frame semantics we take the view that word meanings are best understood in reference to the conceptual structures which support and motivate them. We believe, therefore, that any description of word meanings must begin by identifying such underlying conceptual structures. Frames have many properties of stereotyped scenarios – situations in which speakers expect certain events to occur and states to obtain. In general, frames encode a certain amount of 'real-world knowledge' in schematized form.
>
> (p. 2-3)

The typical conceptual structures include but not limited to *metaphor, metonymy, image schemas, blending,* and *mental spaces*[4]. *Metaphor* and *metonymy* are modelled in terms of mappings: *metaphor* involves mappings between two things from different areas of knowledge while *metonymy* involves mappings between two things within the same area of knowledge. *Image schemas* arise from repeated instances of embodied experience (Hamawand 2016). *Blending*, also known as *conceptual integration*, involves blending two mental spaces to create a new one, the feature of which is not found in either of the input spaces (Hamawand 2016). *Mental spaces* are built up in discourse to provide "a cognitive substrate for reasoning and for interfacing with the world" (Fauconnier 1997: 34). Mental spaces are set up dynamically throughout the ongoing discourse based on linguistic and non-linguistic information (Fauconnier 2012) for interpretative purposes (Hamawand 2016) to realize local understanding and action (Fauconnier and Turner 2006: 307). Fauconnier (1994) treats mental spaces as "partial structures that proliferate when we think and talk, allowing a fine-grained partitioning of our discourse and knowledge structures" (p. 11).

[4] See Hamawand (2016: 102) for a summary of features of the conceptual structures.

Concepts do not occur as isolated in mind but can only be comprehended and retrieved in a context of presupposed background knowledge structures – *domains* (Langacker 1987; Lakoff 1987a). Langacker (1987) defined a domain as "a coherent area of conceptualization relative to which semantic units may be characterized" (p. 488). Theories of domain hold the view that "domains provide a way of carving out the scope of concepts relevant for characterizing the meanings of linguistic units" (Cienki 2012: 14). Domains are perceived as something broader than mental spaces: "whereas mental spaces involve conceptualizations enlisted by the individual in a specific context for a specific purpose, domains encompass many aspects of an experience that are conceptualized as associated" (Cienki 2012: 14).

The introduction and application of *domains* in semantics research has refreshed the traditional views of meaning construction (e.g. the generative view, etc.). The suffering of the latter concerns the limited data scope which may not be sufficient to reveal "complete situations that include highly structured background knowledge" (Fauconnier 1997: 7-8). As meaning construction is negotiated by participants in communication, the on-line creativity is crucial to mediate the relation between "real-word events" and "linguistic facts" (language behaviors) through "elaborate human cognitive constructions and construals" (Fauconnier 1997: 8). What domains do is to provide interlocutors with structural prompts of expectations beyond the immediate communication, which assists "various kinds of reasoning, on-line meaning construction, and negotiation of meaning" (Fauconnier 1997: 8). The conceptual and structural coherence of a domain for knowledge representation helps to make sense of individual linguistic facts and contextual cues in a scenario by logically connecting them to achieve discursive coherence.

Domains do not exist or perform in an isolated manner. A considerable number of studies

have shown that mappings between cognitive domains are set up when people think and talk (e.g. Reddy 1993; Lakoff and Johnson 1980a, 1980b; Turner 1986, 1993; Lakoff 1987b; Norman 1988; Lakoff and Turner 1989; Sweetser 1990; Sweetser and Fauconnier 1996; Cameron and Deignan 2006, etc.). Evidence has been mounting to prove the central role of various kinds of mappings in natural language use. Fauconnier (1997) examined three different kinds of mappings:

- *Projection mappings*: A mapping projects part of the structure of one domain onto another. The general idea is that in order to talk about and think about some domains (called target domains) people use the structure of other domains (called source domains) and the corresponding vocabulary.
- *Pragmatic function mappings*: Two relevant domains (which may be set up locally) typically correspond to two categories of objects, which are mapped onto each other by a pragmatic function.
- *Schema mappings*: This kind of mapping operates when a general schema, frame, or model is used to structure a situation in context. Mappings of such kind involve more fine-grained partitioning of discourse and knowledge structures; namely, mental space building within or across domains.

(p. 9-12)

If a *concept* can be seen as a "mental unit", a *domain* is the structured background knowledge to represent a concept (Clausner and Croft 1999: 3). Langacker (1987) argued that every concept is characterized relative to a semantic domain and that any structured experience has a potential to be treated as a domain if it supports conceptual knowledge related to presupposed knowledge. The concept-domain relationship, according to Clausner and Croft (1999), involves a conceptual transformation process of categorization; specifically, any concept as a categorical member of its relational domain can in turn function as a domain to categorize other concepts. Domains are essential to understand language functioning in

communication. They are central to the interpretation of both semantic and pragmatic meanings of linguistic expressions and to the negotiation of meaning for knowledge construction.

2.3.2 *Frame and framing in motivating contexts*

Fillmore's (1982a) definition of *frame* ("a system of categories structured in accordance with some motivating context", p. 381) highlights the motivating function of *context* to interpret the application of cognitive frames and framing-in-interaction. Frame Semantics holds the view that lexical meanings retrieved from such interpretation is highly context-dependent. Fillmore (1982a) described motivating context as follows:

> The motivating context is some body of understandings, some pattern of practice, or some history of social institutions, against which we find intelligible the creation of a particular category in the history of the language community.
>
> (p. 381)

The situated interpretations always go beyond the boundary of the linguistic structure of the immediate discourse, with "words and their semantic potentials point to, allude to or admit of certain in situ interpretation" (Linell 1998a: 127). There is inherently incompleteness in language use and interpretation and what is perceived by participants to be relevant for meaning construction is based upon their making sense of/with a series of contextual resources. Linell (1998a) defined contextual resources as "potential contexts that can be made into actual, relevant contexts through the activities of the interlocutors in dialogue" (p. 128). He provided a list of contextual resources (original emphases):

1. *Immediate* contextual resources

    1) *Co-text*: the prior discourse in the encounter;

    2) The surrounding *concrete* (material, physical, embodied) *situation* or circumstantial setting, which is spatially and temporally specific;

2. *Mediate* (abstract) contextual resources

    1) What actors already assume, believe, know or understand about what is talked about in the discourse, which is often termed '*model'* of discourse-in-context;

    2) *Current and upcoming communicative projects* (updated through discourse);

    3) Specific knowledge or assumptions about persons involved (actors knowing each other in the discourse);

    4) The *abstract situation definition*, defining encounters as an instance of a certain activity type (see Levinson 1979) or situate activity system (see Goffman 1961) (what is going on in a situation);

    5) The *specific organizational context* (actors knowing each other's role in the discourse against their social roles as background knowledge);

    6) The *socio-historically constituted contexts of institutions and (sub) cultures*;

    7) *Knowledge of language, communicative routines and action types*;

    8) *General background knowledge*.

(p. 128-131)

There are four points I need to emphasize here concerning how the contextual resources are to be properly understood to assist my later discussion of the research findings. First, the contextual resources touch different aspects of how meaning potentials are to be aware of and selected as relevant in the ongoing discourse. Specifically, the model of discourse concerns knowledge of register/genre of discourse; the organizational-communicative projects/routines and knowledge of action types concern knowledge of speech acts and sequential structures of interaction (e.g. turn-taking, topical boundaries, etc.); assumptions about each other's

biographies and discursive roles concern knowledge of identity and participation framework; and general background knowledge embedded in social, historical, institutional contexts concerns knowledge of social connectedness and power relations.

Second, the contextual resources overlap with each other in a "complex matrix of contexts" (Linell 1998a: 132). How the contextual resources are expected to be applied and managed may not be completely shared by participants in a given talk scenario (Linell 1998a), which may lead to interpretative gaps or even misunderstandings in the interaction.

Third, the fourth contextual resource cited above, namely, the *abstract situation definition*, is interchangeably referred to as *frame* by Linell. Note that the "situatedness" of interpretation is emphasized in the meaning of the notion, and a *frame* conceptualized this way "sets up an expectation structure among actors" (Linell 1998a: 130). This explains why some scholars with contextual perspectives towards discourse interpretation (e.g. MacLachlan and Reid 1994) prefer the term *framing* to *frame*, since the former better serves "to reify dynamic processes" (Linell 1998a: 132) of discursive interpretations.

Fourth, it is important to pay attention to the role of socio-institutional factors in terms of their shaping power of discursive interpretation through framings. As emphasized by Linell (1998a), this is of particular relevance to the analysis of "professional discourse" (e.g. academic talk) and "professional-lay interaction"[5] (e.g. consultancy meeting) (p. 131), thus is compatible with the research interest of this study.

[5] Professional-lay discourse is referred by Linell (1998b) to a type of discourse where professionals meet and interact with lay people.

2.3.3 *Frame and framing in social interaction*

Framing is one of the key concepts in the research of social interaction. The notion of *frame* was first introduced into the field of ethnography and ecological studies of society in the 1970s for an approach towards contextual discourse analysis of human social interaction and experience (Tannen 1993). Goffman (1974) proposed the concept of *frame* based on the earlier work on ethnography of communication (see Hymes 1961) to analyze the organization of human experience in moment-to-moment interaction. *Frames* are conceptualized by Goffman (1986) to consist of "principles of organizations which govern events - at least social ones - and our subjective involvement in them" (p. 10). *Framing*, then, is defined as the ways humans in any interactive situations communicate and discern "what is it that's going on here?" (Goffman 1974: 8), "a filtering process through which societal-level values and principles of conduct are transformed and refocused so as to apply to the situation at hand" (Gumperz 2003: 3).

Frame and framing in sociological studies have been widely discussed and seen various definitions emerging from different research contexts. Bateson (1972: 186) defined a frame as "a class or set of messages (or meaningful actions)" which provides cues about how interactants would define an interaction and interpret the communicative acts within the specific context. Gitlin (1980) defined frames as "persistent selection, emphasis, and exclusion" (p. 7) in his study of news media. Gamson and Modigliani (1987) saw a frame as a "central organizing idea or story line that provides meaning" (p. 143). Drake and Donohue (1996) defined frames in a negotiation context as "the particular quality assigned to an issue by the negotiator's linguistic choices" (p. 201). In addition, Putnam and Holmer (1992) conceptualized framing as a communicative process in which interactants create and negotiate

meaning through verbal/nonverbal descriptions and inter-relationship representations. Kendon (1992), adopting an ethnomethodological view, drew upon a "collaborative, multiparty concept of framing" (p. 324) arguing that individual speakers' use of body language and spatial orientation at the action level is worth equal (if not more) attention with framing operations in verbal utterances.

Goffman's (1974) notion of *framing* shares a lot with what is called *interactional frames* in Frame Semantics-"the framing of the actual communication situation" (Fillmore 1982a: 378). This kind of framing deals with how people "conceptualize what is going on" between interlocutors (Fillmore 1982a: 379), concerning an abstract structure of expectations which contributes to defining the roles, purposes, conventionalized sequences of language-in-action associated with certain frames. However, Goffman was not concerned with how grammatical and lexical units would function to frame "what is being said" and "what is conveyed" in the interactional process (Gumperz 2003: 3) and how such units would be used along with para-linguistic mechanisms to reflect and shape context. Therefore, a gap left at both the theoretical and practical levels.

Efforts have been made to integrate the psychological and sociological dimensions of framing to characterize different communicative discourses. For example, Pan and Kosicki (1993) saw framing as "a strategy of constructing and processing news discourse" since it involves applying "cognitive device in information encoding, interpreting, and retrieving" (p. 57). Tannen and Wallat (1993) analyzed the spoken data collected from a pediatric examination and suggested "a particular relationship between interactive frames and knowledge schemas by which a mismatch in schemas triggers a shifting of frames" (p. 61). Carragee and Roefs (2004), by problematizing the existing conceptual approaches to framing in news text

analysis, argued that framing processes need to be examined "within the context of the distribution of political and social power" (p. 214).

Adequate conceptualizations of framing in the previous research have highlighted how framing involves the social construction of meaning. These can be seen as development or correction, as argued by some scholars, of Goffman's notion of *framing*. Re-considering Goffman's *Frame Analysis* (1974), Denzin and Keller (1981) suggested that the work did not provide concepts useful for studying social interaction since Goffman preferred structural models constituting established rules to "freeze interaction into a single frame, into a single answer regarding what is going on" (p. 59). That meaning is embodied in such structural configurations ignores the subjectivity of social interaction and "the multiple realities of the different individuals in the situation" (Denzi and Keller 1981: 59). A more constructivist perspective is suggested to address the interpretive aspect of meaning construction. This general idea has been contextualized in more recent studies which tend to define and remedy the weaknesses in existing framing research (e.g. Carragee and Roefs 2004). A new research agenda has been proposed to conduct framing analysis to address the functional dimension of discourse production and meaning negotiation in social interaction.

The analysis of frame and framing in interaction, according to Telles-Ribeiro and Hoyle (2009: 74), is "a way of studying the organization of experience", "an approach to cognition and interaction that focuses on the construction, conveying and interpretation of meanings". That frames are "interactively achieved" (Telles-Ribeiro and Hoyle 2009: 7; see also Fillmore 1982a; Kendon 1992; Tannen and Wallat 1993) helps to understand what is going on in interaction and ways speakers signal and interpret the interactional activities they are engaged in and jointly constructing.

In his discussion on the analysis of frames in talk, Goffman (1974) proposed three points concerning how language use functions as framing devices in human interaction. First, he argued that the role of words can be a source of both framing and misframing in a conversation for their recipient. The speaker can break frames just as he/she can create and utilize frames through the way he/she manages the production of words. Second, institutional talk has been found to be less loosely inter-connected to surrounding events than in informal conversations. The latter "seems not to be closely geared into extensive social projects, but rather occurs as a means by which the actor handles himself during passing moments" (Goffman 1974: 501). The former, by contrast, tends to show a closer interrelationship between language use and the context it is embedded in. Unlike informal talk at each juncture of which "a whole range of actions seems available to the individual" (Goffman 1974: 501), institutional talk usually allows limited choices of language resources and a heightened use of procedures which would narrow the range of available actions.

Third, ways of framing can be various according to how the speaker chooses to reproduce a scene to his listener(s), or as Goffman (1974) put, "to replay it" (p. 504). This involves the speaker's evaluation of the moment-to-moment interaction as well as his/her intention to conceptualize the talk to his/her listeners, so that the listeners can "empathetically insert themselves into" (Goffman 1974: 504) the talk. This concerns a higher level of shared intentionality in interaction, coined as *joint intention* by Tomasello (1995), referring to more complex forms of cooperation through which individual participants represent the plans of one another and coordinates his/her intentions according to the plans and the overall goals of the communication (Gärdenfors 2014).

The application of framing devices has also been classified and evaluated in other research contexts. Some scholars focused on the structurality of framing practice. Pan and Kosicki (1993) categorized framing devices used in news discourse within four dimensions: syntactical, script, thematic, and rhetorical. Lexical choices in these dimensions are taken as "designators" to "establish a correspondence" between linguistic forms and their cognitive representations (Pan and Kosicki 1993: 62). Other categories of framing devices are found to be "especially pertinent to the sequestering of organizational narratives" (Clair 1993: 118); for example, "accepting the dominant interests as universal", "simple misunderstanding", "reification" (Giddens 1979); "trivialization" (Gitlin 1980); "denotative hesitancy" (Schiappa 2003); "private/public expression and private/public domain" (Sennett 1977).

Much similar to Fillmore's (1982a) categorization, Tannen and Wallat (1993:59) saw the various uses of frame in diverse research contexts falling into two categories: "interactive 'frames of interpretation'" and "knowledge structures which we refer to as 'schemas' of concepts. The former refers to "a sense of what activity being engaged in, how speakers mean what they say" (Tannen and Wallat 1993: 59) and the latter refers to "participants' expectations about people, objects, events and settings" based on their prior knowledge of the world (Tannen and Wallat 1993: 60). Instead of following such a strict categorization, the current study adopted an alternative view that framing has a character of being twofold: the stability of what occurs as meanings which are shaped by speakers' prior knowledge and the variability of interaction which makes meanings continuously negotiated in the context.

The following example offers a glimpse of how the understanding of a lexically evoked frame enters into the sequential analysis of talk-in-interaction. In the example, two speakers ($2 and $4) at a group project meeting are talking about the temperature of a generator that gets too

high.

The overcompensated generator (Marine Engineering, NUCASE)

```
1  <$4>  It ((the temperature)) ↓will(.) become too ↓high
2  <$2>  ↑Yep
3  <$4>  Right
4  <$2>  Eventually
5  <$4>  =So do we almost have to O:VER-(1.6) °What's the ↑word°
6  <$2>  Overcom-
7  <$4>  ↑YEAH(.)Overcompen↑sate
```

The linguistic form *high* in Line 1 evokes the [MEASURABLE_ATTRIBUTES][6] frame. This frame conceptualizes an entity which has a particular scalar attribute with a value exceeding a contextually-specified or generally-understood standard (FrameNet _frame definition). The contextual specificity of the particular standard (the maximum temperature of the generator as allowed by the operating system) is equally accessible to the interlocutors. That $4 knows that $2 shares the understanding of the standard further leads the former to invoke a corresponding frame of [SURPASSING], by which a terminology (*overcompensate*) is referred to (Line 5). The stability of the shared professional knowledge contributes to a joint attention to the terminology, which is achieved by $2's utterance of it in the full (Line 7).

The management of framing at the conceptual level, to this point, has contextual relevance to the management of talk sequences where $4 initiates a clarification of the terminology from $2 (°What's the ↑word°, Line 5) after failing to come up with it himself (O:VER-, Line 5). In addition, $4's management of prosody manifests the variability of interaction in the moment-to-moment negotiation of meaning. Specifically, his utterance of the incomplete form

[6] In this study, small capital letters in square brackets are used to represent frames.

of the terminology with a stretched sound and an increased volume indicates a trouble source of communication. This leads him to place himself lower at the relative epistemic position (see Section 3.3.2 for a detailed discussion), which is revealed from the lower volume of his question. The meaning negotiated this way, therefore, is not only associated with the lexical concepts encoded by the linguistic forms but also attributed to how the state of knowledge can be contextually managed in the interaction.

2.4 Knowledge construction in university small group talk

This section aims to contextualize knowledge construction in the setting of the current study. Previous studies on knowledge construction in collaborative environment were examined and problematized. Theoretical and practical challenges faced by researchers to approach the relationship between knowledge construction and language use in university small group talk are articulated.

The discussion in Section 2.2 suggests that collaboration is an intrinsically important aspect of education at all levels. Interactive talk is a useful tool for education (Berge 1999; Quantz 2001) by which both the teacher and learners are expected to exchange ideas, negotiate meaning and achieve communicative goals in an interactive fashion (DeKlerk 1995). Previous research in discourse studies has covered a broad range of issues concerning collaborative work in educational contexts, such as interactional structures of group talk (Berrill 1991; Coultard 1992; Basturkman 2002), tutor-student interaction (Jungwirth 1993; Viechnicki 1997; Walsh 2002), ‘space for learning’ created through interactive teacher talk (Walsh and Li 2013, 2016), conversation analysis of turn-taking patterns (Markee 1995), and topic development in university seminars (Button and Casey 1988; Stokoe 2000; Benwell and Stokoe 2002; Gibson et al. 2006).

Specifically, Mercer (1995) outlined three essential requirements to guided collaborative knowledge construction:

1. Explain how language is used to create joint knowledge and understanding;
2. Explain how people help other people to learn;
3. Take account of the special nature and purpose of formal education.

(p. 66)

The three requirements highlight the nature of being "shared" "helpful" and "goal-oriented" of knowledge construction in educational practices. The meaning of shared knowledge is "made explicit, clarified and negotiated in an interpretive process" (Arvaja 2007: 135). The shared understanding can further develop understanding of the concepts and procedures about the subject at hand, which would induce more negotiation for learning (Arvaja 2007).

The knowledge collaboratively constructed in such "a cyclic process" (Arvaja 2007: 135) involves continuous exchanges between meanings, ideas, concepts and viewpoints among participants, whose language use is of great importance when the construction of conceptual representations is integrated with specific institutional goals. In academic written discourse, writers must take into account the expected readers' reactions to their work, since it is ultimately the latter that "provide the social justification that transforms beliefs into knowledge" (Rorty et al. 1980: 170). Academic spoken discourse shares this "intertextual framework" for one's "local knowledge" (Berkenkotter and Huckin 1995: 59), within which

interlocutors must consider each other's reaction to and evaluation of what is claimed as valid knowledge.

Small group work as a key component of academic activities has cognitive and motivational benefits to optimize teaching and learning effects (Dillenbourg et al. 1996). Active and productive engagement in group interaction and collaborative meaning production lead to richer cognitive gains (Barron 2003; van Boxtel et al. 2000) and promote reasoning at higher conceptual levels (Amato and Amato 2005). The existing studies have focused on various dimensions of small group work, such as assessment (Gatfield 1999), psychological factors (Cantwell and Andrews 2002), motivational outcomes (Boekaerts and Minnaert 2006), task interdependence (Johnson and Johnson 1999; Gillies 2003), general team atmosphere (Johnson et al. 1998; Wooten and Reed 2000), and cultural homogeneity/heterogeneity (Behfar et al. 2006; Halverson and Tirmizi 2008; Popov et al. 2012).

Specifically, Cutting (2001) conducted a longitudinal study of the conversations of six students as they become members of an academic discourse community, focusing on the speech acts of their language use when involved in group interaction. A model of speech act categories was suggested to take into account factors such as in-group membership and participants' attitude to characterize the talk. The social rules and norms of the in-group were explored from a communicative view and their implications for group solidarity were suggested. Cutting (2015) interviewed 178 German English language teachers and students concerning their awareness of vague language use in academic talk. The finding showed that vague epistemic stance indicators that indexed explicitly a lack of knowledge, as perceived by the participants, might not be appropriate in formal educational settings. The research had implications for both learners and educators to optimize teaching and learning.

The question that to what extent the teacher and learners share a common understanding of the assumption of participation in small group work (Ewald 2008) has been situated at the centre of discussion on its efficacy to facilitate teaching and learning (see Good and Grouws 1975; Brophy and Evertson 1978; Smagorinsky and Fly 1993; Foorman and Torgesen 2001). On the other hand, disagreements arise in terms of an optimal size of a small group (how "small" a group should be to be called "a small group"). Some studies in peer interaction literature suggested that pair-interaction functions better because there is no chance to withdraw or leave others the responsibility to speak (Webb 1989). Other studies argued that larger groups (e.g. of four or five persons) could be more effective to create learning opportunities by encouraging learners to consider a wider range of ideas and communicative resources, thus make the discussion more open and last longer (Needham 1987; Alexopoulou and Driver 1996).

Higher Education has been embracing the benefits of small group talk (Boud et al. 1999; Gupta 2004) to facilitate generic learning outcomes and to promote learners' skills of critical thinking, team co-operation, self-directed learning and interpersonal communication (Kimmel and Volet 2010). On the other hand, university classrooms have been increasingly representing characteristics as a social institution (Atwood et al. 2010) and the talk in its settings has shown more and more features of "talk at work", or "institutional interaction" (Drew and Heritage 1992: 3). Small group talk at the university level has, thus, been taken as a variant of institutional talk in consideration of how meaning is constructed and mutual understanding developed through discursively-shaped talk (Atwood et al. 2010).

The research field has seen some of the most recent studies conducted from the perspective of interaction using different methodologies for data interpretation. For example, Evison (2012, 2013) used Corpus Linguistics techniques to explore multiple turn openings in conjunction with comparative measures of turn-initial priming in spoken academic discourse. The research was aimed to identify what makes academic talk "academic" in respect of role/goal-driven behaviors in interactionally asymmetrical encounters. The findings suggested a greater focus on tutor reflection on roles, goals and discourse in their teaching practice as part of their professional development. Walsh et al. (2011), O'Keeffe and Walsh (2012) and Walsh and Knight (2016) combined Corpus Linguistics and Conversation Analysis to provide an enhanced analysis of spoken interaction in the context of small group teaching in Higher Education. The two approaches were combined in an iterative process to account for features of talk at both micro (word) and macro (text) levels to provide powerful insights into the ways in which participants construct meaning and establish understandings in educational settings.

It is quite challenging, however, to develop theories which specifically focus on "how the goals and conventions generally associated with 'the university' shape talk" (Atwood et al. 2010: 364) in group format. Several reasons are involved to explain the difficulty. First, there is a significant change on almost every aspect of learning in Higher Educational settings. Higher Educational practice explicates changes on knowledge construction, skill development and interactional patterns, most of which involve cognitive and operational capabilities at higher levels. For example, Kaartinen and Kumpulainen (2002) found that university students are more likely than primary and secondary students to comprehend new ideas using prior conceptual understandings and more likely to "draw on their cultural and personal backgrounds as resources for discussion" (Atwood et al. 2010: 363-364).

Second, the relationship between agents in educational activities ("teachers" and "students" as conventionally defined) has undergone a significant change. University students as mature beings have seen a rise of awareness of social identity, which contributes to more complicated social inclinations when engaged in learning activities. University teachers usually act as facilitators and adopt a relatively neutral theoretical position (Atwood et al. 2010) to encourage exchanges of diverse ideas (Scott et al. 2006). For example, a shift has been identified "from the traditional classroom hierarchy towards student resistance to academic identity" in university group work (Benwell and Stokoe 2002: 448).

Third, the industrialization in Higher Education, especially in most Western countries like the UK, has been exerting its shaping power to university students' attitudes towards the teaching-learning relationship and the ultimate purpose of knowledge acquisition. The "teacher-student" relationship in conventional classrooms has been shifted to one of "provider-consumer" of educational services. Students are "shopping between courses and institutions", aiming to make themselves "employable" (Benwell and Stokoe 2002: 449) in the modern society which emphasizes the productivity and marketability of knowledge. This sort of instrumentalism, however, has frequently been taken by many academics and educationalists as representing a crisis of educational values (e.g. Considine 1994; Riesman 1998; Gumport 2000).

Fourth, the concept of "identity" has seen divergent and sometimes conflicting interpretations in university students' group work. This has serious implications for participants to understand how the process of knowledge construction enters into particular institutional practice, which can be inevitably complicated. The complicatedness mainly refers to the relationship between "a shared identity in group work" and "individual identities among

participants". The former makes people a "group" and the latter creates interactional dynamics by drawing upon the diversity in the group. The former defines the most fundamental rules and regulations of behaviors if one wants to be a member of the group, while the latter opens the space for possibilities and variations concerning how the organization of talk-at-work would be shaped.

Finally, the generic features of university small group talk cannot be described in a clear-and-cut fashion. There are different forms of knowledge and different patterns of interaction affecting each other to make the talk unique. Factors that would have an impact on the structures of interaction and knowledge may work in concrete and specific combinations and be open to moment-to-moment adjustment and re-arrangement to achieve particular interpretive purposes (Stivers et al. 2011).

Much earlier work on university small group talk had focused on the relationship between group dynamics and negotiation of knowledge, particularly interested in how groups structure opportunities for individual group members' engagement in knowledge production and construction (Kittleson and Southerland 2004). For example, Richmond and Striley (1996) investigated the process of meaning making by which students solved science problems in small groups. They observed the discourse patterns in four-member groups and explored how different types of leadership would contribute to forming a framework within which social interaction norms are directed and knowledge is co-constructed. Bianchini (1997) investigated how students used the Complex Instruction model to construct scientific knowledge in group work. He identified an unequal pattern of participation in group work where students with higher status, or as defined with "perceived academic ability and popularity" (p. 1042), would contribute more to the group efficacy. The qualitative analyses showed an asymmetrical

power distribution in the groups where students with higher status had greater access to the group resources at material and discursive levels. Kelly and Crawford (1997) investigated the discourse patterns in peer groups and found that there was inequality existing in the availability of learning opportunities among the students and that not all group members had an equal opportunity to construct knowledge through the interaction.

Moje and Shepardson (1998) found that social and institutional factors had significant influences on how participation was structured in group work, which in turn shaped the forms of knowledge construction due to the inequality in the access to the group's academic and interactional resources. Similarly, Anton (1999) investigated learner-centred and teacher-centred discourses in second language classrooms from a sociocultural perspective and found that the former provided more opportunities for meaning negotiation than the latter. The language used to serve the functions of scaffolding was found to effectively assist learners' negotiation of meaning through various linguistic and communicative resources.

More recent studies have paid more attention to the complex combinations of various socio-cultural factors which influence the process and result of knowledge construction in small group settings. For example, Wang et al. (2009) investigated the extent to which individual reflections, group collaborations and class discussions promoted students' critical thinking and knowledge construction in an interactive learning environment. A content analysis was used to analyse 17 college students' reflections and discussions. The findings indicated that not all the students thought critically and that collaboration in group and class discussions contributed to social knowledge construction. Chang (2010), using social constructivism for data interpretation, explored how culture as a tool supported knowledge construction in a learning community in China. The results indicated the multi-dimensional nature of culture to

support knowledge creators to construct new knowledge and to spread new knowledge among knowledge receivers. Bocchi and Cianci (2012) investigated the relationship between contexts and different forms of knowledge production, by reflecting on issues like boundaries between different kinds of knowledge, relevance of knowledge type to contexts, the role of creativity for knowledge construction, etc. Gash (2015) viewed collaborative knowledge construction in education from three interrelated perspectives: how the individual learner builds knowledge, how this construction may be facilitated by collaborative learning using peer-interaction, and the broader institutional context of the educational process in which measures can be taken to develop new ways of supporting collaboration. Mayordomo and Onrubia (2015) explored relationships between the coordinating organization of small group work and the level of knowledge construction in computer-supported collaborative learning environment. The results showed the existence of such a relationship, in which different types of group organizations were related to different levels of knowledge construction and task performance.

The previous research cited above has been overtly focusing on how the organization of group practice would have impacts on its members' performance to construct knowledge. While a broad range of factors (e.g. knowledge types, participatory structures, leaners' reflection and creativity, collaborative formats, etc.) have been addressed in these studies, there was a lack of comprehensive investigation of the very mechanism and process of knowledge construction in university small group talk in terms of:

- What is the most basic unit of meaning for knowledge construction?
- How can the units of meaning be identified and interpreted using valid methods?
- How is knowledge constructed through connection of the units of meaning in specific epistemic domains?
- What is the relationship between knowledge construction forms, discourse structures, and institutional routines?

In the context of university small group settings, less attention has been paid to the semantic-pragmatic interface involved in the actual process of knowledge construction through collaborative talk; namely, how the knowledge represented by the meaning of certain lexical concepts enters into the process of inferential interpretation to achieve mutual understanding among participants. The way conceptual structures are assembled for the purpose of meaning-making requires analyses of talk at the interactional level in terms of how participants' linguistic awareness is raised to facilitate effective communication for knowledge construction (Fung 2011).

2.5 Summary

In this chapter, some of the previous studies on knowledge construction were outlined and discussed, with a particular focus on the research which was conducted in collaborative settings involving spoken language use.

Section 2.2 discussed relevant theories of knowledge construction within three dimensions: the linguistic dimension, the discourse dimension, and the social dimension. Language facilitates the production of meaning in communication by providing the basis for thinking and categorizing life experience into schematized scenarios. The negotiation of meaning plays a vital role in constructing knowledge, relying on the combination of lexical concepts associated with speakers' linguistic choices. Lexical concepts have semantic potentials to provide access sites to conceptual knowledge structures for the interpretation of meaning. The forms of knowledge and structures of discourse mutually shaped each other with the latter providing ways meaning is created through language use. Structural configurations of

discourse are fundamental to understanding meaning-making for knowledge construction through linguistic practices set up across contexts. Knowledge construction involves moment-to-moment expression and negotiation of social relationships in fluid and dynamic processes.

Section 2.3 focused on frames as structures of knowledge construction in semantics research and social interaction research. The theory of Frame Semantics has been discussed after Minsky's notion of frame was introduced and critically evaluated. Frames are regarded as systems of linguistic choices and categories structured in accordance with some motivating context. Frames provide the fundamental representation of knowledge in human cognition and shape relational structures for communication which are flexible and context-dependent. Framing, accordingly, is perceived as ways human in interactive situations communicate to organize experience through moment-to-moment meaning negotiation. An example was given at the end of this section to 1) explain how the notions of frame and framing in Frame Semantics and social interaction can be conceptually related in the current study; and 2) show how the understanding of a frame evoked by a particular lexical concept enters into the analysis of meaning-making and –negotiation at cognitive and communicative levels.

Section 2.4 intensively reviewed some of the empirical studies on knowledge construction in the setting of university small group talk. The research focuses and implications of these studies were discussed. Challenges were identified concerning how the goals and conventions associated with the university would possibly shape small group talk at both the conceptual and the practical levels. A comprehensive investigation of the processes of knowledge construction is needed to focus on the structural characteristics of meaning-making from micro perspectives.

Chapter 3. Methodology

3.1 Introduction

This chapter proposed a combination of Corpus Linguistics and Interactional Linguistics approaches to data analysis in this study. The methodological synergy was aimed at examining knowledge construction through framing at both macro and micro levels. Corpus Linguistics methods provide the distribution and function of linguistic patterns in a descriptive manner to facilitate comparative investigations at the macro level of discourse (Walsh et al. 2011). On the other hand, Interactional Linguistics analyses focus on microscopic details in the sequential organization of framing-in-interaction (Kern and Selting 2013) and are useful to reveal the moment-to-moment negotiation of meaning and construction of knowledge. Both approaches are intrinsically empirical and focus on actual patterns of language in use (Biber et al. 1996; O'Keeffe and Walsh 2012). In addition, complementary grounds can be identified in terms of the data scopes and analytical perspectives for each of the approach, which makes the methodological synergy plausible and accessible.

The purpose of this chapter was to introduce and explain the methodological synergy, the theoretical principles and epistemological underpinnings of each synergetic component. This is necessary to understand the analyses and discussions in subsequent chapters. The chapter is organized as follows: Section 3.2 focused on Corpus Linguistics methods relevant to the study. Following a brief review of previous studies which used corpus-based methods to approach knowledge construction (Section 3.2.1), an overview of the approach (Section 3.2.2) was given. The notion of keyness was introduced to expand the analytical scope beyond the boundary of keywords into key semantic domains (Section 3.2.3). Two computer programs for Corpus Linguistics analyses were introduced with their basic functionalities explained: the

USAS system for semantic tagging of the corpus data (Section 3.2.4) and the WordSmith Tools 6.0 for collocation calculation and concordancing (Section 3.2.5). Section 3.3 focused on Interactional Linguistics methods which are relevant to the study. The approach was briefly overviewed in Section 3.3.1. An intensive review of the ethnomethodology in social interaction was given to address the theoretical and empirical bases upon which the method has been grounded (Section 3.3.2). Section 3.4 provided the rationale (Section 3.4.1) of combining Corpus Linguistics and Interactional Linguistics to analyse the NUCASE data, with the method of the combination outlined (Section 3.4.2). Section 3.5 summarized the whole chapter.

3.2 Corpus Linguistics

3.2.1 *Corpus-based methods to approach knowledge construction*

Some other scholars apply corpus-based methods to investigate knowledge through examinations of lexis use. The approach owes a great deal to the advent and development of computer technology in the field of Applied Linguistics. Compared with manual compiling, computers have made it easier to store and retrieve large collections of language data in electronic form and to analyse them using "increasingly sophisticated, versatile and user-friendly software tools" (Altenberg and Granger 2002: 1). Using a corpus has advantages in lexis study since 1) "lexis lends itself perfectly to the form-based research (e.g. letters, lemmas, word spaces, punctuations, etc.) at which computers excel" (Altenberg and Granger 2002: 1); and 2) the distribution of lexis can be easily computed and observed from frequency counts of words (Hunston 2002), which forms the basis for more complex and sophisticated computation of linguistic relations. These two factors make the task of examining lexis use straightforward (Moon 2010) and the examination inductive rather than intuitive.

Although scholars using corpora to approach lexis do not usually address knowledge construction in an explicit way, their work has serious implications for identifying knowledge structures through patterned lexis use in discourse which might be overlooked without the assistance of computer techniques. Patterns emerging at the lexico-grammatical level (see Halliday 1966; Sinclair 1991; Biber et al. 1998) have particular relevance to the restrictiveness of linguistic choices for meaning-making, the process of which would reveal the scope and preference of knowledge construction through natural language use.

Knowledge structures representing schematized human experience are related to formulaic language use constituting patterned word associations. The linguistic research on word associations has been focusing on the existence of collocations, the combinational restrictions reflecting "the habitual or customary places" of words in company (Firth 1957: 12). Earlier work had paid attention to the computation and attribution of collocational patterns based on marked word senses and uses (e.g. Biber 1993; Smadja 1993; Stubbs 1995; Williams 1998). Numerous empirical studies followed to broaden the investigatory scope by explicating the implications of words' collocational behaviours in generic discourses (e.g. Xiao and McEnery 2006, language learning; McEnery 2006a, 2006b, bad language and moral panic; Baker et al. 2008, media discourse; Siyanova and Schmitt 2008, second language production and processing). More recent research, interestingly, has seen a trend of re-considering and improving the existing methodological procedures to address the multi-dimensionality of collocational phenomena (e.g. Evert 2010; Pecina 2010; Gries 2013; Brezina et al. 2015; Baker 2016).

Closely related to meaning representation through collocations is the research on semantic prosody. Arising from Corpus Linguistics, the term *semantic prosody* was attributed to

Sinclair (1991) and first introduced to the public by Louw (1993). In his search for extended units of meaning using concordance lines, Sinclair (1996) found that many words occur frequently in recurring sequences which reveal textual patterns of meaning-making. Stubbs (2002) referred such observable semantic relations between a given word and its typical collocates to *semantic prosody* (p. 225), "a form of meaning which is established through the proximity of a consistent series of collocates" (Louw 2000: 57). Semantic prosody has been viewed to express speaker/writer attitude or evaluation (Louw 2000; Xiao and McEnery 2006; Bednarek 2008), which contributes to revealing how a collocational structure is to be interpreted functionally (Sinclair 1996). Such investigation may lead the researcher "close to the boundary of the lexical item" (Sinclair 1996: 34) to identify the basic units of meaning. The attitudinal functions of collocations, however, are far from explicit and categorical. While Sinclair (2004) saw semantic prosody as an obligatory property, Partington (2004) regarded it as gradable by drawing a binary distinction between positive and negative attitudinal meanings. This concerns how knowledge represented by collocations is to be constructed across contexts (Whitsitt 2005); namely, to what extent the evaluative knowledge of a frequent collocation found in one context "carries over" to another (Hunston 2002: 141).

Another group of scholars tended to approach the relationship between lexis use and knowledge construction by conducting focused semantic analyses of language data. Many researchers have been working on knowledge-based *Word Sense Disambiguation* (WSD) using a broad range of corpus approaches (Mihalcea 2006). For example, the eXtended WordNet (Mihalcea and Moldovan 2001), large collections of semantic preferences retrieved from SemCor (Agirre and Martinez 2001) and BNC (McCarthy 2001), large scale topic signatures acquired from BNC (Cuadros et al. 2005). Specifically, Cuadros and Rigau (2007) evaluated the relative quality of available knowledge resources on a WSD task to build a large

and rich knowledge base for broad-coverage semantic processing. Hassan et al. (2007) introduced a system to identify lexical substitutions (McCarthy 2002) for words in a given context by combining knowledge sources. Other researchers evaluated the accuracy of knowledge-based approaches to semantic tagging of corpus data (e.g. Andreevskaia and Bergler 2007) and to semantic relations between lexis of a certain part-of-speech in English (e.g. Tribble and Fahlman 2007; Beamer et al. 2007).

Specifically, the notion of *frame* as a knowledge structure (Fillmore 1982a, 1982b) has been introduced and applied to corpus-based lexis processing. The contribution of FrameNet data to practical lexicography and *natural language processing* (NLP) has been extensively discussed (see Atkins et al. 2003; Fillmore et al. 2003; Petruck et al. 2004). Building upon the seminal work, Ruppenhofer et al. (2006) provided a comprehensive introduction to the FrameNet Project concerning how texts in a corpus can be reasonably grouped at the semantic level by identifying frames and frame-to-frame relations based on systematic annotation. Baker et al. (2007) designed a task to recognize words and phrases that evoke semantic frames defined in the FrameNet Project and to explore the semantic dependency between them. Litkowski (2007) integrated and exploited FrameNet data focusing on text processing in a knowledge management system to explore the feasibility of a dictionary-based approach to extraction of frames from a corpus.

As shown above, the researchers using corpus-based methods have overwhelmingly focused on the methodological procedures to identify valid units of meaning at the lexico-grammatical level and to explain language phenomena related to representation and construction of knowledge. In spite of the enhanced scope and reliability of Corpus Linguistics analysis (see Biber et al. 1998), the mechanism and process of knowledge construction through natural

language use have not been sufficiently addressed in the field concerning:

- How corpus technical procedures (e.g. annotation/tagging, frequency, keywords, collocations) can benefit knowledge perspectives to approach natural language use rather than vice versa;
- How words and word clusters identified as linguistic forms encode different lexical concepts to achieve textual coherence in expanded discourse rather than within limited window spans;
- How frames evoked by lexical concepts contribute to characterizing collaborative spoken discourse[7] where knowledge is constructed through interaction.

Section 3.3 and Section 3.4 examined the two approaches involved in the methodological synergy for data analysis: Corpus Linguistics and Interactional Linguistics. An overview of each approach was provided in each section, followed by a detailed discussion on the principles and techniques relevant to the research focus.

3.2.2 *Overview of the Corpus Linguistic approach*

The word "corpus" is Latin for "body" (with the plural form "corpora"). A corpus is defined by Baker et al. (2006: 48) as "a collection of texts stored in an electronic database" containing "thousands or millions of words" of "machine-readable text". A corpus is often (but not always) designed and compiled following some principle, thus the texts are to be selected so that they can be taken as "representative of a particular language variety or genre, therefore acting as a standard reference" (Baker et al. 2006: 48). Sinclair (1995), however, tended to define a corpus in a more flexible way by proposing certain linguistic criteria and characteristics that a corpus is assumed to have: quantity, quality, simplicity, and documentation.

[7] See McCarthy and Carter (1997); O'Keeffe et al. (2007) for investigations on lexis in spoken discourse.

Biber et al. (1998) listed some essential characteristics of Corpus Linguistics research:

1. It is empirical, analyzing the actual patterns of use in natural texts;
2. It utilizes a large and principled collection of natural texts, known as a "corpus', as the basis for analysis;
3. It makes extensive use of computer for analysis, using both automatic and interactive techniques;
4. It depends on both quantitative and qualitative analytical techniques.

(p. 4)

According to Viana et al. (2011), a major strength of corpus analysis is that the counts of linguistics forms which are readily identifiable for a computer can highlight language patterns that may be overlooked or ignored through researcher intuition. Biber et al (1998) listed three strengths of the application of computers in Corpus Linguistics research:

- The use of computers help to enhance the scope and reliability of data;
- It enables linguists to investigate the language used in a "natural" occurring context;
- It allows qualitative interpretation instead of a single approach.

(p. 4)

One of the values of corpus approaches is that it allows researchers to identify and analyze complex "association patterns: the systematic ways in which linguistic features are used in association with other linguistic and non-linguistic features" (Biber et al. 1998: 5). A broad range of research questions can be asked concerning the use of a linguistic feature and the characteristics of texts or language varieties (Biber et al. 1998). In addition, corpus analyses

allow researchers to study co-occurrence of linguistic features to explore "continuous relationships" which mean that patterns are not "absolute statements" of frequency in a corpus (Biber et al. 1998: 8) but can be viewed on a continuum in terms of their distributional and functional features.

Corpus approaches have obvious strengths in investigating language in use by highlighting bottom-up processing methods, as opposed to traditional linguistic studies which tend to exclusively focus on language structures. Corpus analyses can be easily compatible with various quantitative and qualitative techniques and can be applied as a complementary approach to many other approaches in language research and other disciplines as well. In terms of research design, Corpus Linguistics methodology emphasizes total accountability which means that the methodology avoids biased selection and cherry-picking criticism in the sampling procedure and that descriptions on the language use in a corpus account for all the data in the dataset.

Compelling insights have been brought together recently from Corpus Linguistics techniques and cognitive approaches to discourse (Cheung 2009). The answer to the question "why and how language users make specific lexical choices in their communication" requires an amendment to the received view in conventional semantics research (e.g. Cruse 1986, Lexical Semantics) on the nature of corpus data. In spite of the criticism on corpus analyses concerning the risk of de-contextualization, the qualitative examination can assist the researcher to go deep into the data for salient patterns and deviant cases, which is of particular interest in a specialized corpus with a relatively small size. In addition, corpus techniques (e.g. automatic semantic tagging) lend replicability and reliability (Prentice 2010) to research findings, thus has great potential to establish patterns of how certain discourse is organized

(Baker 2006).

3.2.3 *From keywords to key semantic domains*

In Corpus Linguistics, a "keyword" is statistically defined based on the work of British linguist Mike Scott. Scott (1999) derived keywords through specific statistical processes. A word is said to be "key" if:

a. it occurs in the text at least as many times as the user has specified as Minimum Frequency;
b. its frequency in the text when compared with its frequency in a reference corpus is such that the statistical probability as computed by an appropriate procedure is smaller than or equal to a *p* value specified by the user.

Scott (1999)

According to Stubbs (2010: 25), keywords are "words which are significantly more frequent in a sample of text than would be expected, given their frequency in a large general reference corpus". This demonstrates that Scott's work has been largely motivated to reveal language patterns which construct texts in their original context (Rayson 2002). The semantic and functional aspects of words are emphasized and the delicate relationship between words and their actual meanings in context is investigated from an extended stretch of phrases or texts beyond the immediate span of words to the left and right of a node word (Stubbs 2010).

One major merit of Scott's approaches to conceptualize and retrieve keywords is that they provide an empirical method based on distribution and frequency, which contributes to identifying the content distinctiveness of texts by highlighting semantically related keywords in a descriptive manner (Stubbs 2010). However, there has been a paradox that a mismatch is

often involved between what human readers of a text perceive to be keywords and what identified by the statistical procedure (Scott 2000b). This is because the text-based information cannot fully represent the original circumstances of the language event due to the multimodal nature of discourse, let alone the conceptual or psychological factors which cannot be directly observed such as the mood of the discourse producer (Rayson 2002).

Scott's WordSmith Tools allows a frequency list taken from one corpus to be compared with one taken from another corpus. Tribble (2000) described the way WordSmith Tools retrieves keywords:

> 1. Frequency sorted wordlists are generated for a "reference corpus" (a collection that is larger than the individual text or collection of texts which will be studied), and for the research text or texts.
> 2. Each word in the research text is compared with its equivalent in the reference text and the program makes a judgment as to whether or not there is a statistically significant difference between the frequencies of the word in the different corpora. The statistical text evaluates the difference between counts per type and total words in each text and can be based either on a chi-square test for outstandingness or a log-likelihood procedure.
> 3. The wordlist for the research corpus is recorded in terms of the "keyness" of each word.
>
> Tribble (2000: 79-80)

A keyword list generated through such procedures contains two types of keywords: positive keywords (words that appear in the focus corpus more often than in the reference corpus) and negative keywords (words that appear in the focus corpus less often than in the reference corpus), which corresponds to the terms "over-use" and "under-use" based on the comparison (Rayson 2002: 52). Two corresponding keyword lists of a similar length will be generated when two corpora of equal size are compared. When the focus corpus is compared with a

much larger reference corpus, positive keywords will appear on the keyword list, along with a smaller number of negative keywords (Baker 2004). Two thresholds need to be set when computing keywords. One is the cut-off point set by establishing a minimum significance of *p* value (which is 0.000001 by default in WordSmith Tools) based on either a chi-square test or a log-likelihood procedure. The other is a minimum threshold of occurrence for each word in the focus corpus (which is 2 by default in WordSmith Tools), although "this does result in manually identified keywords being omitted from the keyword database" (Scott 2001: 118).

What is even more promising is the work to expand the notion of "keyness" from single lexical words to other units of meaning within various linguistic dimensions. Baker (2004) argued that there is no reason why enquiries of keyness should be limited to single words and that comparisons can be achieved by extending units of meaning and examining key clusters rather than keywords. Mahlberg (2007) also discussed the extension of the keyword approach to clusters to study literary stylistics. Lists of key clusters are particularly useful to help the researcher discover patterns of language use which cannot be obtained by merely looking at single keywords.

The frequency similarity, as shown in relevant columns on keyword lists, does not reveal much itself since a word may have a range of possible meanings and can be used within different semantic fields or to perform different pragmatic functions. Investigation of key clusters, on the other hand, can expand the research scenario beyond the boundary at the descriptive level and help to explore the systematic relationship between form and meaning, by examining key lexical elements of a corpus which are expected to create "a dense network of intercollocation, including both continuous and discontinuous phraseological patterns" (Bondi 2010: 4). Keywords/key clusters are ideal candidates for combining quantitative and

qualitative analyses due to the fact that keyness in its very nature is context-bound and reflects the topicality, style and "aboutness" (Phillips 1989) of the text(s). Keywords/key clusters are useful "signposts" to direct analysis from the immediate context surrounding the target words/clusters to expanded text, in which way it portrays a "model of discourse" (McEnery 2015) to highlight features of language use in specific contexts, written or spoken.

In addition, enquiries of keyness can be conducted within various linguistic dimensions. In his work of corpora comparison, Rayson (2008) extended the notion of keyness from single lexical words to other linguistic domains by applying the keyness calculation on grammatical (part-of-speech, also known as POS) and semantic domains of the corpora. He argued that the combination of the keywords and key domains methods allows "macroscopic analysis (the study of the characteristics of whole texts or varieties of language) to inform the microscopic level (focusing on the use of a particular linguistic feature)" for further investigation (Rayson 2008: 519). The main contribution of Rayson's work can be perceived at different levels. At the practical level, the method extends the keyness technique thus provides a way to combine corpus-based and corpus-driven approaches at data analysis and interpretation stages. At the theoretical level, it contributes to a variable perspective that description of language use in a text cannot simply be sufficient within a single domain. If differences can be observed, presented and interpreted across linguistic domains, it may raise the issue of complexity as a result of inter-domain relationships since language use never involves random choices (Kilgarriff 2005).

3.2.4 *USAS: the semantic tagging system*

The application of computer software in Corpus Linguistics studies enables the researcher to re-arrange the way natural language use can be processed thus makes it easily and

systematically accessible (Hunston 2002; Scott 2010). Tagging or annotation is the practice of adding interpretative, linguistic information to a corpus (Leech 1997: 2). There are many reasons for a corpus to be tagged (see Leech and Smith 1999), some of which are summarized by Rayson (2002):

1. Adding further annotations: word-class tagging is a useful first step and simplifies the tasks of syntactic annotation (parsing), semantic annotation, discourse annotation (see below for examples);
2. Information extraction: extracting frequency information, lemmatisation, and collocations from corpora;
3. Information retrieval: document filtering dependent on content;
4. Word processing: spelling and grammar checkers;
5. Speech processing: synthesis and recognition;
6. Handwriting recognition: language modeling;
7. Machine-aided translation: annotation of multi-lingual corpora;
8. Dictionaries and grammars: discriminating homographs for lexicographers writing corpus-based dictionaries;
9. Language learning: students examining real data for grammatical structures used by native speakers;
10. Development of NLP software: training corpus for a part-of-speech tagger or evaluation corpus for a parser;

(p. 17-18)

Tagging of a corpus can be conducted at multiple levels: orthographic, phonological, phonetic or phonemic, morphological, lemma, prosodic, grammatical (POS), syntactic, semantic, discoursal, pragmatic, stylistic, problem-oriented, etc. (see Rayson 2002: 19-21 for details of features and examples at these levels). In this study, semantic field annotation (Rayson and Wilson 1996) was applied to tag both the focus corpus (NUCASE) and the reference corpus (MICASE). Semantic tagging is found to be useful to process patterns of language use in the focus corpus in terms of the following aspects:

- Examining words collected into semantic categories allows the researcher to see trends of language use which are often invisible at the word level (Rayson 2002: 112; see also Henry and Roseberry 2001: 101).
- Not only single words, but also multi-word units are counted together within certain semantic categories. These bundles of lexical or functional words (Biber 2009) are "formulaic sequences" which appear to be "pre-fabricated" (Wray and Perkins 2000: 1) and "pre-constructed" to constitute single choices (Sinclair 1991: 110). The formulaicity of such multi-word units are conceptually related to the application of cognitive frames in communication, since both of them concern how knowledge is stored, retrieved and constructed through patterned language use.
- Variants of a lemma are usually grouped together into a semantic category (Rayson 2002), which is useful to investigate the relationship between word meaning and its syntactic structure.
- Since the semantic tagging is designed to apply to open-class or 'content' words, it is particularly useful to identify semantic categories consisting of disciplinary lexical concepts represented by nouns or noun phrases. This helps the researcher locate her analytical focus accurately and quickly in a sizable dataset.

Software-based automated semantic tagging, as one of the widely applied Corpus Linguistics techniques, facilitates enquires on knowledge representation and construction by identifying lexical items within unique semantic categories in the target discourse. The automated semantic tagging system used in this study is the UCREL Semantic Analysis System (USAS)[8] developed by Rayson (2002) and available at the Wmatrix[9] interface. The UCREL semantic analysis system is a framework for undertaking the automatic semantic analysis of a text. It accepts the text which has been tagged for parts of speech and feeds it into the main semantic analysis program (SEMTAG) which assigns semantic tags representing the general sense fields of words from a lexicon of single words and a list of multi-word expressions.

[8] The semantic tagset used by USAS was originally loosely based on Tom McArthur's *Longman Lexicon of Contemporary English* (McArthur 1981).

[9] Wmatrix is a tool for statistical analysis of corpora developed by Paul Rayson at Lancaster University, UK.

Currently, the lexicon contains nearly 37,000 words and the template list contains over 16,000 multi-word expressions (Rayson 2002). It has a multi-tier structure with 21 major discourse fields, subdivided into a further 232 sub-categories, and with the possibility of further fine-grained subdivision in certain cases. Typically, the semantic tags are composed of:

- an upper case letter indicating general discourse field.
- a digit indicating a first subdivision of the field.
- (optionally) a decimal point followed by a further digit to indicate a finer subdivision.
- (optionally) one or more 'pluses' or 'minuses' to indicate a positive or negative position on a semantic scale.
- (optionally) a slash followed by a second tag to indicate clear double membership of categories.
- (optionally) a left square bracket followed by 'i' to indicate a semantic template (multi-word unit).

Rayson (2002: 66)

The authors of the USAS tagging system (see Appendix A) combine various techniques on semantic tagging and WSD, classify items according to a semantic field taxonomy rather than definitions of word meaning, and assign semantic categories to all words rather than selected classes (Rayson et al. 2004). The multiple-source WSD approach helps to achieve a high degree of tagging accuracy (Rayson et al. 2004): a 91.05 percent precision rate has been obtained from the tested and refined USAS system, leaving an error rate of 8.95 percent (Rayson et al. 2004). Figure 4 shows what a semantically tagged text looks like in the USAS system.

```
0000001 002  -----   -----
0000004 001  NULL    <$1>
0000004 010  PPH1    It                  Z8
0000004 020  VVZ     depends             A2.2 S7.1-
0000004 030  RRQ     how                 Z5 A13.3
0000004 040  PPIS2   we                  Z8
0000004 050  VV0     look                X3.4[i1.2.1 X2.4[i1.2.1 X2.1[i1.2.1 A8 X3.4 X2.4 X7+ Z4
0000004 060  II      at                  X3.4[i1.2.2 X2.4[i1.2.2 X2.1[i1.2.2 Z5
0000004 070  AT      the                 Z5
0000004 080  NN1     chairman            S7.1+/S2mf
0000004 081  .       .
0000004 082  -----   -----
0000006 001  NULL    <$2>
0000006 010  UH      Yeah                Z4
0000006 020  RR      exactly             A4.2+ A5.3+
0000006 030  UH      yeah                Z4
0000006 031  .       .
0000006 032  -----   -----
0000006 040  CS      If                  Z7
0000006 050  AT      the                 Z5
0000006 060  NN1     chairman            S7.1+/S2mf
0000006 070  VBZ     is                  A3+ Z5
0000006 080  RR      merely              A13.6 A14
0000006 090  AT1     a                   Z5
0000006 100  NN1     person              S2mfc
0000006 110  PNQS    who                 Z8
0000006 120  VVZ     runs                M1/N3.8+ G1.2 G2.1- X3.3 S7.1+ A1.1.1 T2++ M3 S4 A6.3+ K4
0000006 130  AT      the                 Z5
0000006 140  NN2     meetings            S1.1.3+ S3.1 K5.1 S9 A1.1.1 X9.2+ A6.1+
0000006 141  .       .
```

Figure 4 Sample of USAS tagged text in vertical format

When effectively used with keyword procedures, semantic tagging helps to identify key semantic domains. Rayson (2002) presented a case study on a dataset of field reports of a series of ethnographic studies at an air traffic control center (ATC). The focus corpus consisted of verbatim transcripts of the ethnographer's observations and interviews with controllers and of reports for later analysis by a multi-disciplinary team of social scientists and system engineers. A 2.3 million-word subset of the BNC including the transcripts of spoken English was used as the reference corpus. Both the focus corpus and the reference corpus were tagged using the semantic field tagger in the USAS system. Log Likelihood test was applied as the statistical metric to calculate the keyness value. The comparison represented a list of key semantic tags which were deviant from the reference corpus. Nearly all (except for one) of the key semantic categories were found to include important objects, roles, functions, etc. in the ATC domain, thus were particularly useful to characterize the specification of the discourse and uncover the concerns of the discourse producer(s).

The same method was used in this study to do semantic tagging and generate key semantic domain lists through comparisons between the focus corpus and the reference corpus. The semantic comparison reduced the number of key categories the researcher would further examine for fine-grained analysis (Rayson 2002). This method is totally data-driven based on frequency values of linguistic forms and statistical tests which inform empirical analysis at both the macroscopic (e.g. characteristics or variations of a target discourse) and microscopic (e.g. the use of a particular linguistic feature) levels (Rayson 2002: 155). The technique also allows straightforward replication of the analytical procedure, thus, ensures the reliability of the research findings (see Prentice 2010).

3.2.5 *WordSmith Tools: concordance lines and collocational patterns*

Another corpus analysis software used in this study is WordSmith Tools 6.0 developed by Mike Scott. It is an integrated suite of programs for exploring how linguistic items behave in texts. It has three main tools for corpus analysis: *Concord* is used to create concordances in the language data, enabling functions of searching and sorting according to the position of target linguistic items; *WordList* lists all the words (or word clusters) included in the focus corpus along with a series of statistical data, which can be sorted by frequency or in alphabetical order. *Keywords* creates a list of keywords (or word clusters) based on the comparison between two corpora. Statistical tests are run in the keywords procedure to identify keywords with unique frequency in the focus corpus compared with that in the reference corpus at a statistically significant level. The statistical metrics and criteria can be set by the user.

Application of WordSmith Tools 6.0 in this study involves the functions under the *Concord Tool*: *Concordance* and *Collocation*. A concordance is a list of all the occurrences of a

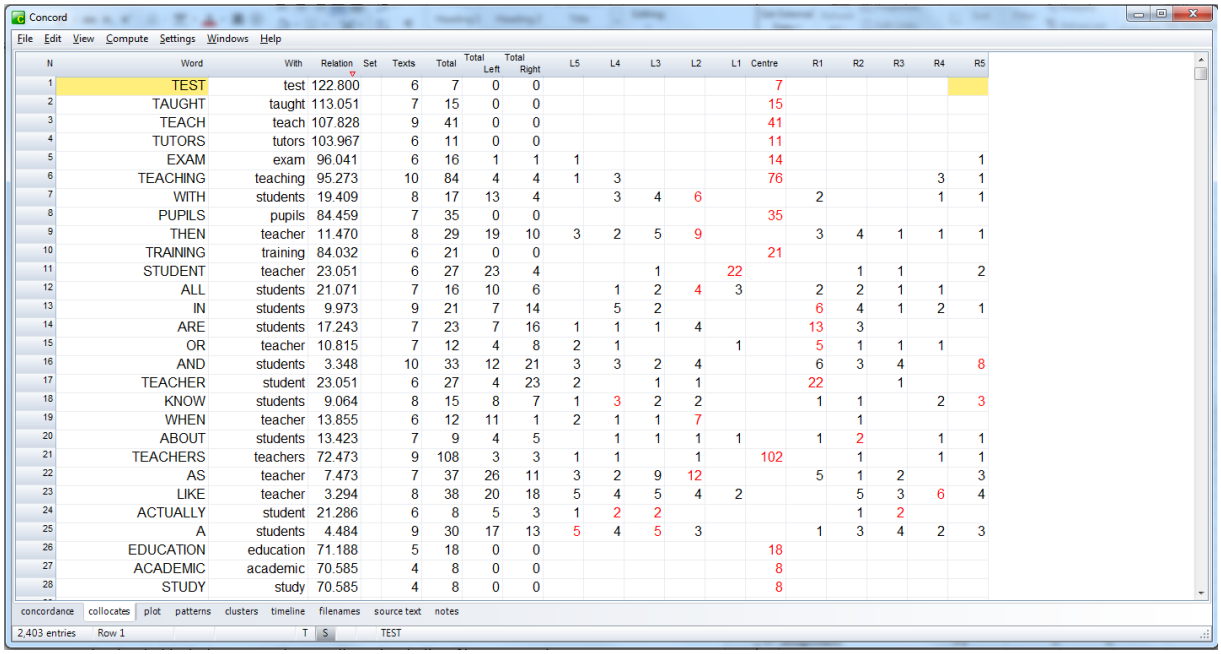
particular search item in a corpus, presented within a limited span (the number of words to the left and the right of the search item) of the context. According to Baker (2006), the object of analyzing concordance lines is to identify patterns of language use based on repetitions and deviant cases in target discourse. Concordance lines have a visual advantage for language processing in a sizable dataset since the subtlety of some instances of language can be difficult to be uncovered by human intuition yet observable "only when a lot of evidence is seen together so that the pattern emerges" (Hunston 2002: 12). *Keywords*[10] *in Context* (KWIC) analysis through concordance lines is one of the most effective CL techniques which allows a closer examination of discourse (Baker 2006). In addition, WordSmith Tools 6.0 has a high functionality which allows the researcher to sort the concordance lines in various ways and can show a set of basic statistics of the search item, including its collocates and plot dispersions. Another merit of the tool is that it can deal with data saved in separate files thus can show the frequency of a search item in each file, which helps to demonstrate its local distribution and to provide information on metadata in data presentation.

In this study, the *Concordance* function is used with another function *Collocation*. The linguistic research on word associations has well developed in the past decades following Firth's (1957) suggestion to look at the "company that words keep" (p. 6). Corpus Linguistics research has been implementing and improving tools to approach collocations in a number of contexts (e.g. Xiao and McEnery 2006; Baker et al. 2008). The research findings have important implications of how a particular field of discourse is organized into observable and accessible lexical patterns (Phillips 1983, 1985, 1989).

[10] "key word" in KWIC has a different meaning from what is called "keyword" in the keyword procedure. The former simply means the word in focus while the latter is defined statistically by comparative frequency.

Collocational relationship is a complex one and cannot be fully addressed with one single association measure[11] (Brezina et al. 2015), -"individual association measures differ in how much emphasis they put on the different criteria" (p. 144). Brezina et al. (2015) reported on the results of a series analyses in a case study on swearing in English (see McEnery 2006a, 2006b) using five different collocation metrics and discusses their efficacy in uncovering feature of the target discourse within different dimensions.

The association measures can be set in the *Concord* setting labeled as *relation statistic*. The researcher can also set other parameters for collocates calculation, e.g. *horizons*, *minimum frequency,* and *breaks*. The horizons represent the number of collocates found to the left and right of the search word. The minimum frequency specifies how frequently collocates must have appeared in the neighborhood of the search word. Breaks limit the scope beyond which collocate calculation shall be stopped (Figure 5 shows an example).



N	Word	With	Relation	Set	Texts	Total	Total Left	Total Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	TEST	test	122.800		6	7	0	0						7					
2	TAUGHT	taught	113.051		7	15	0	0						15					
3	TEACH	teach	107.828		9	41	0	0						41					
4	TUTORS	tutors	103.967		6	11	0	0						11					
5	EXAM	exam	96.041		6	16	1	1	1					14					1
6	TEACHING	teaching	95.273		10	84	4	4	1	3				76				3	1
7	WITH	students	19.409		8	17	13	4		3	4	6			2			1	1
8	PUPILS	pupils	84.459		7	35	0	0						35					
9	THEN	teacher	11.470		8	29	19	10	3	2	5	9			3	4	1	1	1
10	TRAINING	training	84.032		6	21	0	0						21					
11	STUDENT	teacher	23.051		6	27	23	4			1		22			1	1		2
12	ALL	students	21.071		7	16	10	6		1	2	4	3		2	2	1	1	
13	IN	students	9.973		9	21	7	14		5	2				6	4	1	2	1
14	ARE	students	17.243		7	23	7	16	1	1	1	4			13	3			
15	OR	teacher	10.815		7	12	4	8	2	1			1		5	1	1	1	
16	AND	students	3.348		10	33	12	21	3	3	2	4			6	3	4		8
17	TEACHER	student	23.051		6	27	4	23	2		1	1			22		1		
18	KNOW	students	9.064		8	15	8	7	1	3	2	2			1	1		2	3
19	WHEN	teacher	13.855		6	12	11	1	2	1	1	7				1			
20	ABOUT	students	13.423		7	9	4	5		1	1	1	1		1	2		1	1
21	TEACHERS	teachers	72.473		9	108	3	3	1	1		1		102		1		1	1
22	AS	teacher	7.473		7	37	26	11	3	2	9	12			5	1	2		3
23	LIKE	teacher	3.294		8	38	20	18	5	4	5	4	2			5	3	6	4
24	ACTUALLY	student	21.286		6	8	5	3	1	2	2					1	2		
25	A	students	4.484		9	30	17	13	5	4	5	3			1	3	4	2	3
26	EDUCATION	education	71.188		5	18	0	0						18					
27	ACADEMIC	academic	70.585		4	8	0	0						8					
28	STUDY	study	70.585		4	8	0	0						8					

Figure 5 Collocates of search lexical (NUCASE)

[11] See Brezina et al. (2015: 140-141) for the proposed criteria to identify collocations.

The first column *Word* shows the collocates. The second column *With* shows the search-words. The third column *Relation* shows the collocational strength between the search-word and the collocate calculated by the set statistical metric (Log Likelihood in this case). The fifth column *Texts* show the number of texts the collocation appears. The numbers shown starting from the sixth column *Total* are:

- the total number of times the word was found in the neighbourhood of the search word
- the total number of times it came to the left of the search-word
- the total number of times it came to the right of the search-word
- a set of individual frequencies to the left of the search word (5L, i.e. 5 words to the left, 4L .. 1L)

(WordSmith Tools Manual 2015)

3.3 Interactional Linguistics

3.3.1 *Overview of the Interactional Linguistics approach*

Interactional Linguistics (IL) is an integrated approach to Discourse Analysis (DA) to study language in use, aiming at a systematic description of both the forms and functions of linguistic phenomenon in talk-in-interaction (Gumperz 2003; Jaspers 2013; Pan 2013; Kern and Selting 2013). Broadly speaking, it applies to the linguistically orientated research of interactional phenomenon (see reviews by Barth-Weingarten 2008; Couper-Kuhlen and Selting 2001; Hakulinen and Selting 2005; Ono and Couper-Kuhlen 2007; and Selting 2008). It has its origins "in the search for replicable methods of qualitative analysis that account for

our ability to interpret what participants intend to convey in everyday communicative practice" (Gumperz 2003: 1).

Meaning-making and intention-inferring are the basic focuses of IL research to investigate how interactants rely on contextualized knowledge at the discursive level to make themselves heard and understood (Gumperz 2003). IL analysts take a point of view that face-to-face interaction is the engagement in "an ongoing process of negotiation, both to infer what others intend to convey and to monitor how one's own contributions are received" (Gumperz 2003: 4). The basic assumptions of IL can be summarized as follows:

1. Language use is "fundamentally context-sensitive" and "practice-, action-, sequence- and recipient-oriented". Language is conceived of as "an ongoing, emergent product in a social semiotic event" and language is always used in particular tasks and for particular purposes (Kern and Selting 2013: 1).
2. Linguistic phenomenon should be analyzed with regard to the "conversational actions" for which they are deployed by participants and the "sequences they are embedded in" (Kern and Selting 2013).
3. Information about contextual frames is communicated as an important part of interactional processes, which should be fully aware of as to assess participants' intentions. However, sequential analysis cannot by itself account for "situated interpretation" and what is analytically problematic is to "discover how interpretive assessments relate to the linguistic signaling processes through which they are negotiated" (Gumperz 2003: 4).
4. Participants always rely on "local or context-specific background knowledge" to interpret and assess others' communicative intentions. There are conversational inferences existing in the whole interpretive procedure and assessments of communicative intentions taking the form of hypotheses can happen at any one point in an exchange (Gumperz 2003: 5).

IL research adopts a strictly empirical method which is grounded in databases of naturally occurring talk-in-interaction (Kern and Selting 2013). IL analysis in its very nature treats speaking as "a reflexive process" where speech as the product can be seen as "either directly reacting to preceding talk, reflecting a set of immediate circumstances, or responding to past

events, whether directly experienced or indirectly transmitted" (Gumperz 2003: 7). IL methods have been widely used within professional and institutional domains where a wide range of topics can be investigated, such as conversational styles, mismatches and misinterpretations, language/knowledge competence at workplaces, social inequality, discrimination and mutual stereotyping (Pan 2013).

IL approaches discourse at a microscopic level, adopting an insider view to study tensions between the "customary ways" of speaking and the "unpredictability of interaction" within larger social processes (Jaspers 2013: 141). In IL research, linguistic forms and constructions are viewed as "the result of an on-going, real-time process of coordination and interaction between the participants, by-products of their construction of practices and actions in conversation" (Kern and Selting 2013: 2). This means that linguistic analyses using IL as research methodology view language use as a context-sensitive, responsive and emergent product to achieve certain communicative goals in interaction.

IL methods draw mainly on principles of sequential analysis (Kern and Selting 2013: 1) to reveal the dynamic relationship between language use, context, and social practice. From the IL perspectives, sequential analysis "describes just one of the many indexical processes that affect inferencing other than accounting for situated interpretation by itself (Gumperz 2003: 4). There is an ongoing process in which participants' communicative intentions are constantly assessed and hypotheses of such intentions are "either confirmed or rejected in the course of the exchange" (Gumperz 2003).

This study integrates the cognitive and communicative dimensions of framing to approach the participants' interaction in small group talk. It started with a tentative characterization of the

target discourse by identifying what specific frames are evoked by the lexical concepts encoded by particular linguistic forms. It then examined emergent patterns of framing which are conceptually invoked for situated meaning interpretation and interactionally manipulated through the talk sequences. Following the basic procedure for IL investigations, the data analysis in this study was conducted in the following steps:

1. Taking an initial look at the NUCASE data to provide insight into the communicative ecology of the university small group talk in the corpus (Section 3.4);
2. Identifying the key epistemic domains mentioned by the participants in NUCASE through a frequency-based comparison of their language use (Section 4.3.2 and 4.3.3);
3. Finding out through observation the language patterns around target lexical concepts in certain epistemic domains (e.g. on top of the frequency list) for meaning-making, drawing upon the participants' own interpretations of the discourse (Section 5.2 and 5.3);
4. Selecting expanded interactional scenarios where the target lexical concepts are embedded for detailed, contextualized analyses of how knowledge is constructed through framing invoked by the lexical concepts in the interactional sequences (Section 5.4).

3.3.2 *Ethnomethodology to approach knowledge construction*

The perspective on verbal communication adopted by Interactional Linguists is believed to be grounded in the earlier work on ethnomethodology. The application of IL approaches has generated empirical investigations of real situations of speaking which constitute units of social interaction at the action level, not just through structuralist abstractions (Gumperz 2003). Seen as a reflexive process, communicative practice is expected to involve knowledge that goes beyond interlocutors' mental grammar. What such knowledge is and how it would affect meaning-making and -understanding at talk are at the centre of IL research.

Ethnomethodology reject the claim that knowledge constitutes pre-existing social structures

(see Garfinkel 1967) and consider knowledge to be managed through collaboration and governed by social norms (Drew 1991). Scholars follow this research trend argue that epistemic status is "an inherently relative and relational concept" concerning the access to some domain of knowledge (Heritage 2012: 4). The nature of access to such knowledge can be the object of complex sequential manipulation in human interaction (Heritage and Raymond 2005). From such a perspective, the actual possession of particular knowledge is only part of knowledge construction process at the action level, since the relativity of epistemic status also embraces how the knowledge is known and whether interlocutors have the rights and responsibilities to know it (Drew 1991; Sacks 1992; Maynard 2003; Terasaki 2004). Epistemic stance, by contrast, concerns how the relative relationships in terms of knowledge representation, construction and transmission can be expressed moment-by-moment through turn design at talk (Heritage 2012: 6).

Ethnomethodological approaches favour a close, detailed analysis of the recording and transcript of the naturally occurring conversation in actual context as "intelligible socially organized interaction" (Heritage 1984: 1). As an empirical discipline, ethnomethodology allows regularities and patterns of talk to directly emerge from the data with no theoretical constructs or pre-identified conceptual categories imposed over the real orientations of speakers under observation (Liddicoat 2007).

Studies drawing the spirits of ethnomethodology have been conducted in various research contexts to approach knowledge construction through small group work in educational settings. The role of participation has been emphasized in language-mediated learning practice (e.g. Duff 2002) with a broad range of external factors such as learning situations, activities, etc. (e.g. Roschelle and Teasley 2012) taken into account. Group collaboration to

construct knowledge has been examined from the view of representation realized by computer technology, for example, linguistic representation (Suthers 2005), diagrammatic (graphical) representation (Suthers 2005), and pictorial representation (Näykki, and Järvelä 2008) to visualize the conceptual mediation of individual and shared ideas in groups. The situatedness of knowledge construction has been investigated to explore the relationship between the level of knowledge construction and effectiveness of meaning negotiation (Evnitskaya and Morton 2011), learner level (Atwood et al. 2010), institutional identity (Evnitskaya and Morton 2011), and facilitative teacher talk (Atwood et al. 2010). These studies showed that knowledge elements are socially negotiated and emphasized the situatedness of participation in the process. They conducted data-driven qualitative analyses focusing on the sequential structures of talk-in-interaction to reveal features of knowledge construction.

Ethnomethodological approaches have methodological advantages when the researcher focuses on how action formation in human interaction is realized through adjustment of lexicon and prosody. The adjustment is based on assumptions and predictions of the degree of divergence between what speaker A knows and what speaker B knows in an interaction. In other words, the ethnomethodological perspectives are can be particularly useful when speaker A is remarkably more (or less) knowledgeable than speaker B at a particular moment of their interaction (see Labov and Fanshel 1977, 'A-event statement' and 'B-event statement'). Heritage (2012) argues that the degree of equality of access to knowledge varies upon "an epistemic gradient" (p. 4), a slope from being shallow to deep (Heritage 2010; Heritage and Raymond 2012). An epistemic gradient has important implications for understanding the form-function relationship revealed by particular utterances to form certain actions; for example, how questioning is done by non-questioning forms (Schegloff 1984), and how certain types of negative interrogatives are actually used to assert rather than request

information (Bolinger 1957; Koshik 2005; Heinemann 2006). Therefore, it is argued by scholars in this field that status of knowledge is fundamentally relevant to the formation of action and management of interaction (Heritage 2012).

As an integrated approach to discourse analysis, Interactional Linguistics specifically focuses on the interactional aspects of discourse production and understanding. Jaspers (2013) summarized four points concerning why IL is important in discourse analysis, especially in the analysis of talk-in-interaction:

> 1. IL is important because it emphasizes the existence of subtle cultural differences in the systematic combination of verbal and non-verbal signs which signal contexts and construct meaning, differences that are often hard to pin down by those who use them.
> 2. IL illustrates that technically differing conversational styles do not necessarily lead to miscommunications. It invites researchers to look beyond the actual interactional setting and observe how interactants approach and evaluate one another as they are socially positioned and mutually perceived as different beings.
> 3. IL shows that communication is intrinsically a social happening where identities and relations matter and that both the participants' behaviors and perceptions should be interpreted within wider social patterns.
> 4. IL offers an excellent tool for analyzing the tension between here-and-now interaction and more established discursive practices.
>
> (p. 143-144)

On the other hand, although Interactional Linguistics is strongly influenced by Conversation Analysis (CA) at the methodological level (e.g. sequential analysis), it is conceptually distinct from CA. Interactional Linguistics takes an interdisciplinary approach to linguistic analyses while it is (for most cases) rather difficult for CA to be combined with research methods inspired by a social constructionist view (e.g. Corpus Linguistics). One of the aims of IL analysis is to provide a functional description of linguistic structures at collaborative talk

(Kern and Selting 2013), while what is valued in CA is how interlocutors manage communicative resources moment-to-moment.

3.4 Towards a methodological synergy

An initial observation of the NUCASE data reveals distinct features of the corpus which are taken as the basis for the selection of research methods in this study. The university small group talk in NUCASE is found to be:

- naturally occurring language data in actual contexts with no researcher intervention nor instructional processes for data collection;
- (the data) not collected on the basis of random assignment of participants aimed to create valid and meaningful comparisons;
- in high-quality audio-recorded format and broadly transcribed for all small group talk sessions;
- with little pedagogical orientation or consideration;
- with relatively equal access to particular epistemic domains (disciplinary knowledge shared among participants), thus the "epistemic gradient" (Heritage 2012) is expected to be "flatter" than what may be found in mundane conversation;
- with the relatively limited scope of pragmatic adjustment to negotiate epistemic status, with individual roles in group talk assigned and attended according to the task procedure

Based on the research questions, the data analysis in this study requires a methodological approach which

- can identify lexical concepts encoded by linguistic forms across different epistemic domains and what frames are evoked to characterize the epistemic discourse;
- can characterize language patterns emerging around target lexical concepts and how the participants invoke interpretive frames for situated meaning-making; and
- can examine in detail how the participants interactively frame their talk to construct knowledge across contexts.

I, therefore, suggest a methodological synergy to approach knowledge construction in this study: the combination of Corpus Linguistics and Interactional Linguistics analyses. Both of the methods are empirical and rely on data-driven observations. Corpus Linguistics analysis uses computers to identify language patterns emerging from a large amount of data, thus is efficient to characterize the systematic ways in which linguistic forms are organized to evoke particular cognitive frames for discourse characterization. Investigation of semantic tendencies of lexical concepts encoded by the linguistics forms, in addition, can reveal evidence of how the participants invoke interpretive frames for situated meaning construction. Interactional Linguistics analysis, on the other hand, is aimed to conduct linguistically orientated research of interactional phenomenon by focusing on the sequential organization of talk-in-interaction. It deals with the participants' conceptualization and expectation of communicative situations in moment-to-moment interaction thus can reveal how interactive frames are employed for meaning negotiation and knowledge construction.

3.4.1 *Rationale*

The methodological synergy used in this study for data analysis can be seen as an alternative view to the conventionally proposed dichotomy of language conceptualization. The dichotomy polarizes language use into two contrastive fashions: the system/structure perspective vs. the practice/communication perspective. The former focuses on abstract or decontextualized meanings of linguistic forms while the latter emphasizes the relevance of communicative meanings and functions to contexts. Such distinction suffers over-simplification since language use is a complex phenomenon which requires systematic shifts between perspectives. A structural-functional perspective is thus relevant to this study since "language and discourse practices are interdependent" (Linell 2008: 5).

The structural-functional perspective enhances a sense of the relationship between individual mental representations of knowledge and framing in group interaction (Arvaja et al. 2007). The relationship has been explained by the proposition that group interaction stimulates elaboration of knowledge thus facilitates individual cognitive gains (van Boxtel et al. 2000). Linguistic forms are the basis to organize patterns of associations by encoding lexical concepts across conceptual domains. The relational, dynamic and flexible structures of such patterns evoke frames which "provide the fundamental representation of knowledge in human cognition" (Barsalou 1992: 21).

The combination of Corpus Linguistics and Interactional Linguistics analyses in this study concerns how frames as conceptual knowledge structures in discourse organization can be approached methodologically rather than being observed in an ad hoc manner. Experimental results obtained from studies of classic Lexical Semantics do not necessarily indicate anything explicit about how conceptual representations can be properly observed through the instrument (Komatsu 1992). In such research environments, the difficulty of deciding whether a particular observation is valid lies in the ambiguity concerning how the structure (Anderson 1978), function (Komatsu 1992) or process (Armstrong et al. 1983; Lakoff 1987b; Smith and Medin 1981) of the interventional operation on the research target could possibly affect its representation. While experimental results "do specify at a minimum what information must be available to people about instances of a concept" (Komatus 1992: 501), Corpus Linguistics analysis provides empirical evidence from a frequency perspective concerning what other information is relevant and necessary to the conceptualization process. Legitimate sources of data (though not easy to be obtained directly) are getting readily accessible, including knowledge of situated circumstances of language production and speakers' verbal/non-verbal

interactional behaviors concerning discursive inferences and predictions (Fauconnier 1997).

However, Corpus Linguistic analyses can be limited in that they present numbers rather than instances of language use if the researcher starts his/her enquiry from frequency and keyword lists (Viana et al. 2011). Quantitative analyses carried out this way tend to be based on counts of linguistic forms without distinguishing their meanings in use, which would risk total de-contextualization in data interpretation (Viana et al. 2011). Observations of instances of linguistic forms through concordance lines and their collocational patterns are thus necessary to address the qualitative aspect of corpus data. Such analyses, though attending to the surrounding context of a given linguistic form, do not usually go further into expanded discourse beyond a limited word span. It is from here that Interactional Linguistics analyses become relevant and necessary to push the observation beyond the scope of distributive language patterns by carrying out close, detailed sequential analyses of interaction. In such way, emergent framing patterns are to be examined in the on-going interaction to reveal the dynamic relationship between language use, context, and socio-institutional practice. The combination of Corpus Linguistics and Interactional Linguistics analyses take into account situational perspectives as well as computational techniques (see Ross 1970). The methodological synergy not only enhances the internal validity of the data analysis but also helps triangulate the research findings.

3.4.2 *Method*

While the main techniques for analysis featuring a combination of Corpus Linguistics and Discourse Analysis methods now appear to be well-established (e.g. concordances, collocates, keywords, dispersion, etc.), applying corpus methods in discourse analysis, however, is still a developing and a "cross-disciplinary field which is somewhat under-subscribed" (Baker 2006:

6). Using corpora to approach discourse has been subject to as much resistance and criticism as interest during the past decades. Corpus-based approaches have been problematized by some researchers as "too broad" to "facilitate close readings of texts" (cited in Baker 2006: 7) and only offering "a partial account of language" by ignoring the correspondence between corpus findings and native speaker intuition (Widdowson 2000: 7). Other opponents claim that corpus linguistics treats language as a self-contained object by "abstracting text from its context" (Baldry 2000: 36) and that it is often difficult to draw conclusions about language use when an example does not appear in the corpus or is implicitly endowed with meaning by native speakers (Borsley and Ingham 2002).

However, that "all methods of research have associated problems which need to be addressed" (Baker 2006: 7) does not mean that corpus approaches cannot contribute to uncovering interesting things in discourse studies. An epistemological shift to post-structuralism as an alternative way to produce and process knowledge (Denzin 1988) "favors a more eclectic approach to research" (Denzin 1998: 16) in which different methodologies can be combined within localized research context. Baker (2006) summarized the advantages of corpus approaches to discourse analysis as follows:

1. Reducing researcher bias;
2. Revealing the incremental effect of discourse by examining repeated patterns shaped by cognitive and cultural factors;
3. Approaching language dynamics by examining resistant and changing discourses;
4. Achieving triangulation by applying multiple methods of analysis.

Baker (2006: 17)

Nevertheless, not much attention in the literature has been paid to combining different methodologies to approach knowledge construction in group work. A few exceptions (e.g. Hmelo-Silver 2003; Ebenezer et al. 2010) used experimental interventions and statistical tests in combination with more fine-grained analyses. The qualitative analyses, however, are only taken as a complementary approach to the basic experimental designs, or a way of methodological triangulation for more valid results.

In this study, the key semantic domains are directly extracted from the NUCASE data on a frequency basis. The Corpus Linguistics analyses are strictly empirical starting from a comparison between the semantically tagged focus corpus and reference corpus to identify the key epistemic domains. Lexical concepts encoded by specific linguistic forms with high frequency values in the over-used, non-shared key epistemic domains (across the sub corpora) are then investigated through concordance lines and collocations to show patterns of meaning-making for interpretive frames invoking. The lexical concepts are grouped into domains of knowledge with clear epistemic boundaries. This is a way of "slicing" the corpus data (Walsh 2014) and viewing it within categories of meaning rather than by individual words, which assists the analysis of situated meaning-making through more identifiable language patterns. A qualitative analytical method is then employed to zoom into the micro-contexts beyond the statistical results in exploring the relationship between language use, meaning construction and interaction (Christodoulidou 2011). Detailed observations of framing-in-interaction can be carried out via close, contextualized analyses of expanded talk scenarios to show how knowledge is constructed through the participants' manipulation of the interactional sequences.

The nature of "keyness" which has been discussed earlier is of particular relevance to the

proposed methodological synergy. “Keyness” of a linguistic form is represented in the form of a value which is calculated based on the statistically significant difference in its frequencies when two corpora are compared with each other. The larger the keyness value of a linguistic form is, the more key it is in the focus corpus when compared with the reference corpus. This is the quantitative facet of keyness. On the other hand, “keyness” is a textual matter which is highly context-dependent, which means that the reference corpus and the comparison itself matter in the keyword procedure. This is the qualitative facet of keyness which emphasizes the importance of taking a closer look at keywords in context. The researcher needs to be aware that a keyword/key domain with similar keyness values calculated from different comparisons can reveal different discourse patterns and underpin differently situated contexts. The context-dependency of keyness from CL perspectives and the functions of framing devices provide an interface where these two methodologies for data analysis can be mixed at both the conceptual and methodological levels.

The major strengths of using a methodological synergy to approach knowledge construction in the current study are (adapted from Dörnyei 2007) as follows:

- It develops a comprehensive and complete interpretation of knowledge construction in the research context (university small group talk) by increasing the strengths of both methods while eliminating their weaknesses (e.g. de-contextualization of CL, unrepresentativeness of IL);
- It conducts multi-level analyses of complex issues of knowledge construction concerning its nature (e.g. information exchange through meaning negotiation) and representation (e.g. framing through talk of lexical concepts, epistemic status through sequential manipulation);
- It improves the research validity with a potential to produce evidence through the convergence of the findings;
- It attracts multiple audiences with various research interests and paradigmatic orientations, thus, contributes to communication and cooperation between researchers working in different fields.

(p. 45-46)

In sum, the methodological synergy proposed in this chapter demonstrated that the current study is, at the methodological level, complementary to recent studies that have combined Corpus Linguistics and Discourse Analysis approaches to interactive university talk.

3.5 Summary

This chapter introduced and presented an overview of the methodology which was adopted in this study. Section 3.2 reviewed relevant theories and methods of Corpus Linguistics, focusing on the expanded scope of investigation of keyness. The USAS system for semantic tagging and the WordSmith Tools 6.0 for collocation generation and concordancing were introduced. Section 3.3 reviewed relevant theories and methods of Interactional Linguistics. The basic procedure of using the method for sequential analyses of the small group talk was outlined. Based on the reviews, Section 3.4 provided the rationale and method of combining Corpus Linguistics and Interactional Linguistics analyses in this study.

Chapter 4. Research Design

4.1 Introduction

This chapter addressed issues of the research design of the current study. Brief corpus descriptions were given to justify the use of NUCASE data (Section 4.2.1), the sub corpora construction (Section 4.2.2) and the reference corpus selection (Section 4.2.3) for the keyword procedure. NUCASE data was first cleaned to be compatible with the processing format of Wmatrix (Section 4.3.1), upon which key semantic domains were identified based on the grammatically and semantically tagged transcripts (Section 4.3.2). A sampling procedure was imposed to select data for further transcribing and analysis (Section 4.3.3 and 4.3.4). The validity, reliability of the research design and the generalizability of expected results were discussed at the end of the chapter (Section 4.4).

4.2 Corpus description

4.2.1 *NUCASE as a specialized spoken corpus*

The dataset used in the current study is part of a one-million-word corpus: *Newcastle University Corpus of Academic Spoken English* (NUCASE). NUCASE comprises approximately 120 hours of audio-and video-recorded small group talk sessions around Newcastle University. The data cover a broad range of speech events, including seminars, tutorials, PhD supervision meetings, staff-student consultations, language education classes, students' project meetings and presentations, and sessions involving informal learner talk. Learners at both undergraduate and postgraduate levels are included in small group sessions and the participants are fully aware of the relevant ethical issues concerning the corpus construction and the data collection process. Initiated in 2010 at Newcastle University, the corpus is still under construction to meet the various research needs based on the NUCASE

project.

The rationale of using corpus data to investigate features of knowledge construction across academic disciplines lies, to a large extent, in the fact that the "robustness" of the link between knowledge, language, and discipline presented in previous studies cannot be clarified, accessed or represented in a principled way (Nishina 2010). NUCASE is a specialized spoken corpus whose features can be represented according to Flowerdew's (2004) parameters in terms of its specificity (Table 5).

Parameter	NUCASE specificity (relevant to this study)
Purpose for compilation	To identify both linguistic features and interactional resources used by speakers in an academic context; to describe, characterize and evaluate of speakers classroom interactional competence; to demonstrate how certain features are criterial to learner competence levels
Contextualization	Setting: small group sessions in university classrooms Participants: university students and teachers/tutors Communicative purpose: teaching and learning
Size	1,722,451 tokens (584,218 tokens used in this study)
Genre	Academic spoken discourse
Type of text/discourse	University small group talk
Subject matter/topic	Applied Linguistics and Education; Marine Engineering; Bioinformatics; Business and E-business; Computer Science; Pre- and In-sessional English language classes (the former three used in this study)
Variety of English	Mainly learner[12] English

Table 5 Specificity of NUCASE

As O'Keeffe (2007) argued, specialized corpora have a clear advantage over large general ones in terms of providing insights into the context of language use and help the researcher find certain marginal cases which he or she is particularly interested in or has not expected in

advance. Specialized corpora are particularly useful to explore language patterns in the following ways:

1. The patterns are localized in particular context due to the link between their distributional concentration and local situational conditions;

2. The patterns are linked to pragmatically specialized language use in certain context, or the coverage of particular language patterns can be narrowed and concentrated by the specificity of their representations which can lead the researcher to the focus on particular features of language use in given contexts.

In addition, specialized corpora have other advantages over general corpora from a methodological perspective:

1. The analyst is probably also the compiler of his/her specialized corpus/corpora, which means that he/she would be familiar with the wider socio-cultural dimension where the discourse was created (Flowerdew 2003). This position can provide a "mediating ethnographic" perspective which would shed light on the corpus data.

2. More qualitative-based analyses (rather than pure quantitative-based analyses) tend to be carried out on specialized corpora since their size and composition make it more manageable to conduct qualitative analyses. Close examination of concordance lines and detailed analyses on collocational features can provide a richer source of data to reveal language patterns which might not be easy to stand out when observed through the quantitative lens.

3. It is possible and manageable to conduct detailed tagging procedures on specialized corpora so that more top-down and contextually-informed analyses can be done on the semantic and discourse level of the corpus data.

4. Many investigations on specialized corpora are comparative in nature which can "uncover differences almost regardless of size" (Sinclair 2001: xii).

(Flowerdew 2004: 16-18)

4.2.2 *Sub corpora construction*

47 small group talk sessions which last approximately 63 hours are used. The total tokens are 584,218. Investigations are conducted concerning how knowledge is constructed across disciplines. Three sub corpora are constructed according to three academic disciplines: *Bioinformatics*, *Marine Engineering*, and *Applied Linguistics & Education*. The construction work is based on the assumption that knowledge in university small group talk may be constructed in different ways across different academic disciplines. Table 6 shows the basics of the sub corpora:

Main corpus	**NUCASE**			**Total No.**
Sub corpus	Marine Engineering (ME)	Bio-informatics (BI)	Applied Linguistics & Education (ALE)	3
No. of sessions	16	21	10	47
Time duration	20:46:17	27:55:17	13:52:22	62:55:56
Total tokens	287,653	137,147	159,418	584,218

Table 6 NUCASE sub corpora construction

The number of participants in each small group session ranges from 4 to 12 and the time duration of a single session ranges from 15 minutes to 5 hours. Specifically, the sub corpus of Marine Engineering sees all the sessions conducted in the form of student project meetings where teachers or tutors were not involved. The sub corpus of Applied Linguistics & Education includes small group sessions with teacher/tutor and student participation, e.g. tutored teacher training feedback sessions, discussion groups of data analysis. The sub corpus of Bio-informatics consists of student project meetings and consultation meetings held between university student research teams and pharmaceutical companies.

4.2.3 *Reference corpus selection*

Different from comparing individual parts within a corpus, a comparison between corpora always requires the researcher to choose a reference corpus. The selection work is essential since different reference corpora may give different results of keyword analyses. In his search of a bad reference corpus, Scott (2009) investigated a wide variety of factors which may influence the quality of a reference corpus and the results of keyword analyses, which include size, the extent of randomness of selection, strangeness, and relation to the corpus in question, language and genre. He argued that the keyword analysis procedure to identify a text's "aboutness and structuring" will clearly achieve results "which are largely dependent on the qualities of the reference corpus itself" and that comparisons of different reference corpora are necessary for different research purposes (Scott 2009: 80). He further made three general claims concerning the relationship between reference corpus selection and research results:

> 1. That the choice of reference corpus will affect the results;
> 2. That features which are similar in the reference corpus and the node text itself will not surface in the comparison; but
> 3. Only features where there is a significant departure from the reference corpus norm will become prominent for inspection.
>
> Scott (2009: 80)

However, Scott (2009) did not identify a "bad" reference corpus. His findings showed no clear threshold below which poor keyword results can be expected. The quality of the keywords remains unknown when a different reference corpus is used since the same keywords may be generated, whose qualities cannot be easily evaluated on the surface. He

concluded that the keyword procedure is suggested to be "robust" and that "the aboutness of a text may not be one thing but numerous different ones" (Scott 2009: 91).

The relationship between reference corpus selection and research results might be more delicate than ever expected. This poses a question of how to *use* a reference corpus *appropriately* rather than *choosing* an *appropriate* reference corpus to achieve the research aim. In reality, it might be difficult to find a perfectly-matched reference corpus for certain comparison and the researcher needs to justify his or her choice in any sense that the comparison(s) can be made legitimately. According to Scott (2009), one needs to take into account the size, the date, the aboutness, and whether a single text or a whole set of texts is compared with a reference corpus.

If the selection of a reference corpus could possibly affect research results in keyword procedures, both "similarity" and "difference" can be relative rather than absolute concepts between any two corpora and comparisons at various levels can be justified. According to Kilgarriff (2009: 1), keyword procedures tend to work best "where the corpora are very well matched in all regards except the one in question". This provides the rationale for choosing *Michigan Corpus of Academic Spoken English* (MICASE)[12] as the reference corpus in this study. MICASE is one of the few publicly available and web-searchable corpus containing spoken discourse within educational institutions. Table 7 shows the basic features of the two corpora.

[12] See http://quod.lib.umich.edu/m/micase/ for more details of MICASE.

Corpus	NUCASE-relevant to this study (Newcastle University Corpus of Academic Spoken English)	MICASE (Michigan Corpus of Academic Spoken English)
Year of construction	From 2010 until now (still under construction)	1997-2001 (released 2002; new interface 2007)
Corpus type	Corpus of academic spoken English	Corpus of academic spoken English
Corpus size	584,218 tokens, around 62.5 hours	Over 1.8 million (1,848,364) tokens, almost 200 hours
Corpus component (speech event)	47 Small Group Talk sessions (4-12 speakers), including seminars, tutorials, staff-student consultations and informal learner talk across Newcastle University	152 recordings from a wide range of speech events, including lectures, classroom discussions, lab sections, seminars and advising sessions and other locations across Michigan University
Academic discipline	Three academic disciplines: Marine Engineering, Applied Linguistics & Education and Bioinformatics	Various academic disciplines, including Social Sciences, Natural Sciences, Economics, Arts and Humanities
Learner level	Undergraduate and postgraduate	Undergraduate and postgraduate
Speaker status	Native and non-native speakers involved	Native and non-native speakers involved
Interactivity	Highly interactive	Mixed: highly interactive – mostly interactive – mostly monologic – highly monologic
Data collection site	Newcastle University, Three faculties: Humanities and Social Sciences, Medical Sciences and Sciences, Agriculture and & Engineering.	Michigan University, academic events in the professional schools (i.e., medical, dental, business law) were excluded; construction parameters taken into account include interactivity of sessions, learner level, faculty level, native vs non-native speakers, participant gender, etc.
Data format	Video- and audio-recordings	Audio-recordings
Transcript format (original)	Broad transcripts: header to describe context, speaker marked, not semantically or grammatically tagged, not marked with detailed interactional features (overlapping, intonation, emphasis, pause time, etc. no multimodal features)	Broad transcripts: header to describe context, speaker marked, not semantically or grammatically tagged (transcripts in XML tagged format available online), marked with pauses, laugh, inaudible background conversation, no marked with other detailed interactional features
Texts for analysis (with tagging)	POS and semantic tagging (by Wmatrix)	POS and semantic tagging (by Wmatrix)

Table 7 NUCASE vs. MICASE

Table 7 demonstrates that it is appropriate to choose MICASE as the reference corpus in the current study despite the fact that the corpus mainly comprises American English (for NUCASE, the dominant language variety is British English). The similarities in pedagogical goals (teaching in university settings) and types of participants (teachers/tutors and university students) have made the two corpora comparable with variables well controlled according to the research focus.

4.3 Data analysis procedure

The data analysis in this study has been conducted in the following procedure:

1. Construct sub corpora and save the texts in separate files;
2. Merge the files in the sub corpora and the reference corpus, do initial data cleaning;
3. Upload the merged texts into Wmatrix, save and name the folders;
4. Do POS and semantic tagging on the texts using Wmatrix tag wizard;
5. Run keyword procedures based on the semantic tags (each sub corpus vs. reference corpus);
6. Obtain key semantic tag lists for each sub corpus;
7. Select a sample for further analysis (key semantic domains and lexical items);
8. Select target lexical concepts in the key domains;
9. Search and analyse the selected lexical concepts in concordance lines;
10. Select, further transcribe and analyse stretched texts mapped by the lexical concepts;

4.3.1 *Cleaning NUCASE data*

NUCASE data has been broadly transcribed using the same transcribing convention for the *Cambridge and Nottingham Corpus of Discourse in English* (CANCODE)[13]. Following are the conventions used and the explanations on the conventions:

[13] The full transcription conventions used in CANCODE can be found in Fung and Cater (2007).

- Punctuation: . a full stop indicates a completed intonation unit;
  ? a question mark indicates a question;
- Speaker codes: <$1>, <$2>, <$3>, etc.
- Interrupted utterance: +;
- Unfinished word/single letter: =;
- Unfinished sentence, repeat or false start: <$=>… </$=>;
- Transcriber's guess: <$H>…<$/H>;
- Unintelligible/inaudible utterance: <$G>…</$G>;
- Overlapping utterance: <$OL>…</$OL>;
- Transcriber's Comments: <$E>…</$E>.

The transcription conventions produce tags similar to *Extensible Mark-up Language* (XML) tags to indicate what the data are. However, Wmatrix does not count word frequency tagged by XML tags. Uploading data containing transcripts tagged with angled brackets will cause Wmartix to think that XML tags are used in the data and thus not to count the tagged text for word frequency. Since the retrieval of key semantic domains is based on frequency counts of single words and multi-word expressions, keeping the original transcription conventions will risk losing data thus skew the keyword results. Therefore, NUCASE data must be cleaned before it is uploaded into Wmatrix. For the keyword procedures conducted in this study, it is what the speakers actually say that matters to the frequency counting, while the transcriber's comments or any extra-linguistic information are irrelevant (while such information may be relevant for the qualitative analysis). In addition, unintelligible utterances are invalid since they are not transcribed. Therefore, I kept the two tags "<$E>…</$E>" and "<$G>…</$G>" and excluded the tags with angled brackets except for those for the speaker codes. This will lead Wmatrix to ignore any text transcribed by the two conventions thus not to count the word frequency inside; and at the same time, to count the word frequency in the text which is originally tagged by the other conventions.

Another issue is that Wmatrix cannot identify smart (or curly) quotes (‘or’ or “”) automatically, which will lead to misidentification of some linguistic forms thus will skew their frequency counts. For example, it may, by mistake, identify “teacher’s” as “teachers”, which will generate fewer frequency counts of the former and more of the latter than the actual uses. As mentioned before, the only text type compatible with Wmatrix is plain text (.txt). This text type is plain ASCII which uses straight quotes (' or ") rather than smart quotes, while NUCASE data was transcribed using smart quotes even though the original saving format is the plain text type. To maximize the quality of data to be uploaded to Wmatrix, I further cleaned the data through the following steps:

1. Copying and pasting the original text from .txt files to MS-Word files;
2. Saving the text as plain text by choosing the “Save as” function in the “File” menu;
3. Clicking ‘insert line breaks’ and ‘allow character substitution’ in the dialogue box;

The last option will replace Windows apostrophes with ASCII equivalents so that Wmatrix can identify them in the tagging procedures.

4.3.2 *Key semantic domain identification*

In this section, I describe the keyword methods used to compare the sub corpora in NUCASE with MICASE at the semantic level. Several practical issues need to be clarified before the keyword procedures are run for data analysis. First, according to Rayson (2008), it is important that two corpora for comparison do not overlap or that one is not a sub corpus of the other due to independence assumptions. This point has been fully considered in the sub-corpora construction process to make sure that they do not overlap with each other and that

comparisons will not be made across the two sets of sub corpora. Second, all the separate text files (transcript of each small group talk session) in each sub corpus are concatenated[14] into one text file to fit the upload setting of Wmatrix.

The text files in each sub corpus are then merged into one text file with a new name. The sub corpora are then uploaded to the Tag Wizard in Wmatrix for POS and semantic tagging. The same procedures are used to concatenate the files in MICASE and upload it. All the sub corpora and the reference corpus are listed in "My Folder". Clicking on the folder image will direct the researcher to the page for corpus analysis (Figure 6).

Figure 6 Upload the corpus into Tag Wizard in Wmatrix

[14] In the Windows operational system, the concatenation can be realized by typing the command *"copy *.txt newfile.txt"* in the Microsoft DOS command prompt in the directory where all the separate files are stored.

As shown in Figure 7, frequency lists are available to be seen, sorted and saved from the page. A list is also provided on the right of the page from which the researcher can choose his/her reference corpus to run the keyword procedure:

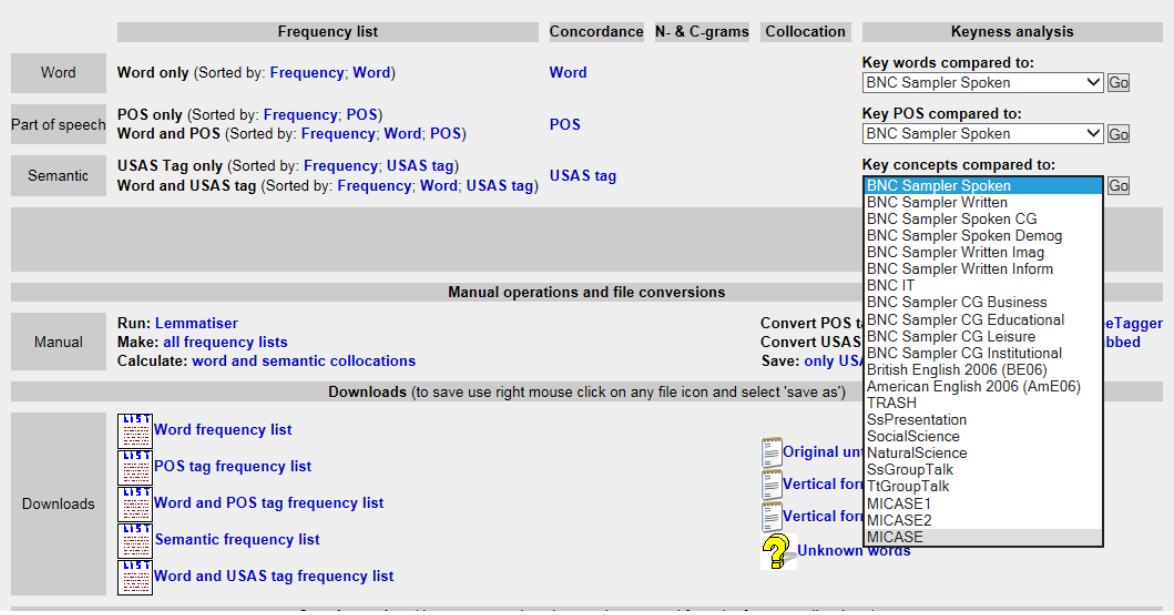


Figure 7 Choosing the reference corpus for the keyword procedure

For each entry in the word frequency lists when two corpora are compared, Wmatrix calculates the log-likelihood (LL) statistic as a statistical significance measure. Note that there is no need to normalize the figures to apply the formula since the calculation of the expected values takes into account the size of the two corpora (Rayson 2008). The higher the log-likelihood value, the more significant the difference is between two frequency scores of a lexical item in the two corpora.

In this study, the log-likelihood value is used in the keyword procedures on semantic domains in the two corpora for comparison. By clicking on the "Go" button to the right of the "Key concepts compared to" bar, the researcher can get a list of key semantic domains which are

used statistically more frequently in the focus corpus than in the reference corpus. Table 8 shows the top 20 key semantic domains in the Applied Linguistics and Education (ALE) sub corpus compared with MICASE:

				Item	O1	%1	O2	%2		LL	LogRatio	
1	List1	List2	Concordance	P1	2109	1.41	16329	0.48	+	1643.65	1.55	Education in general
2	List1	List2	Concordance	Z8	26828	17.92	499076	14.71	+	937.93	0.29	Pronouns
3	List1	List2	Concordance	E4.1+	457	0.31	1548	0.05	+	873.24	2.74	Happy
4	List1	List2	Concordance	Z4	9021	6.03	152983	4.51	+	659.02	0.42	Discourse Bin
5	List1	List2	Concordance	Q2.1	2526	1.69	32756	0.97	+	625.83	0.81	Speech: Communicative
6	List1	List2	Concordance	E6+	116	0.08	362	0.01	+	235.57	2.86	Confident
7	List1	List2	Concordance	X2.1	1648	1.10	25570	0.75	+	199.48	0.55	Thought, belief
8	List1	List2	Concordance	S1.1.2+	210	0.14	1552	0.05	+	175.64	1.62	Reciprocal
9	List1	List2	Concordance	Z6	2869	1.92	50387	1.48	+	163.99	0.37	Negative
10	List1	List2	Concordance	A13.5	270	0.18	2448	0.07	+	160.76	1.32	Degree: Compromisers
11	List1	List2	Concordance	A5.4+	524	0.35	6342	0.19	+	160.27	0.91	Evaluation: Authentic
12	List1	List2	Concordance	X3.2-	73	0.05	244	0.01	+	140.91	2.76	Sound: Quiet
13	List1	List2	Concordance	A14	1278	0.85	20464	0.60	+	131.42	0.50	Exclusivizers/particularizers
14	List1	List2	Concordance	X2.2+	1407	0.94	23019	0.68	+	128.26	0.47	Knowledgeable
15	List1	List2	Concordance	X2.3+	179	0.12	1487	0.04	+	124.48	1.45	Learning
16	List1	List2	Concordance	S1.1.3+	134	0.09	936	0.03	+	121.54	1.70	Participating
17	List1	List2	Concordance	X3.2	228	0.15	2372	0.07	+	102.38	1.12	Sensory: Sound
18	List1	List2	Concordance	X8+	348	0.23	4378	0.13	+	94.90	0.85	Trying hard
19	List1	List2	Concordance	N4	1539	1.03	26860	0.79	+	92.15	0.38	Linear order
20	List1	List2	Concordance	A7+	2649	1.77	49406	1.46	+	90.29	0.28	Likely

Table 8 Top 20 key semantic domains (ALE, LL=6.63)

The O1 column represents the observed frequency of the semantic domain (in the Item column) in the NS sub corpus and the O2 column represents the observed frequency of the same semantic domain in MICASE. 1% and 2% values show relative frequencies in the texts of the two corpora. "+" indicates overuse in O1 relative to O2, while "-" indicates underuse in O1 relative to O2. The list is sorted on log-likelihood value to show the items with the largest keyness values at the top. The "List" links on the left show a list of words grouped with the target tag from the files in the two corpora. The "Concordance" links show the concordance lines of the target tag in the focus corpus. In this way, one can get a list of key semantic domains the concordance lines for all the lexical items for each domain in each sub corpus. The next stage involves my sample selection and examination of concordance lines for analyses on the key semantic domains.

4.3.3 *Sampling procedure*

One of the problems of CL methodology is the variation which cannot be predicted in advance or effectively controlled by the researcher. Not only precision and recall from a corpus can vary, cut-off points and methods of analysis can vary as well within different research contexts. However "representative" or "unbiased" corpus approaches are expected to be as advocated Corpus linguists, they cannot totally remove cherry-picking, since sampling procedures are always involved in analysis and such procedures can be affected by a range of factors such as relevance to research questions, familiarity/unfamiliarity of the language pattern to the researcher, similarity to other language patterns, etc. (Baker 2015).

The relationship between corpus size and representativeness, however, tends to raise more problems when the focus of CL research has begun to shift from using large, general corpora to building and exploring much smaller ones for specific purposes (e.g. NUCASE). A specialized corpus is often small and tends "only to be representative for certain high-frequency linguistic features" (McEnery and Andrew 2001: 78). This means that achieving representativeness can be even trickier since the sampling frame, or the entire population from which the samples would be taken is correspondingly narrowed in size while variations within the whole population might be intensified on the other hand. Maximizing representativeness in specialized corpora can be therefore taken as a task of deriving generalization of specificity.

A close examination is important as noted by many CL scholars that simply getting the significant results is not enough and that it is the researcher's responsibility to explain the results based on the deeper analysis (Woods et al. 1986; Leech and Fallon 1992; Granger 1993; Kretzschmar et al. 1997). The main task at this stage involves selecting a sample from the key semantic domains, identifying the linguistic forms in each semantic domain and

categorizing them in a principled way. Some technical issues are attended to concerning how a sample of target linguistic forms can be selected for analysis.

First, limitations in certain corpus analysis software make the sampling work challenging. For example, the keyword list retrieved through Wmatrix only shows the observed frequencies of a keyword in the focus corpus and the reference corpus, but not how many files where it occurs in the two corpora respectively. The limitation exists because the software was designed to require that the texts in a corpus to be uploaded to be concatenated into one .txt file. WordSmith Tools, on the other hand, can deal with texts in separate .txt files thus includes in its keyword lists the number of texts where the target keyword is used. However, WordSmith Tools cannot tag texts automatically and, even though it can deal with tagged texts, the tagging syntax it reads is not always compatible with what Wmatrix uses for POS/semantic tagging. This means that WordSmith Tools may not be able to identify the tags produced by Wmatrix thus cannot assist analysis on such tags.

In terms of frequency counting, Wmatrix counts single word frequency in a different way from how other software may do. Wmatrix counts multi-word expressions (MWEs) identified by the semantic tagger as one item. It means that if a word appears in an MWE, it will not be counted towards its single word count (Rayson 2016), which would lead to the result that the frequency value of a single search word extracted from the semantically tagged text in Wmatrix might be lower than the one generated by other corpus software (e.g. WordSmith Tools). On the other hand, the tagging syntax would make a difference concerning how the software counts word frequency. Wmatrix would automatically ignore the text tagged with XML tags thus not count the words in the text, while WordSmith Tools would retrieve all occurrences of a search word regardless whether it is XML tagged. This is another reason

why one may expect Wmatrix to work out a different frequency value of a single search word compared with WordSmith Tools if XML tagging is involved.

Second, the key semantic domains rather than the lexical items in the domains seem to be legitimate units of analysis for keyword analysis, since the latter are not necessarily "key" in the corpus. In the keyword procedure, what are involved in the comparison are the semantic domains labelling the lexical items rather than the lexical items on their own. This means that the lexical items are identified because the extra-textual information (semantic senses) added to them makes the frequency of "a semantic sense group" statistically significant when compared with the reference corpus. In other words, individual lexical items, whose frequencies are not directly taken into the keyword calculation, are not keywords at all by definition and their status of keyness is unclear without comparisons at the word level. That is why Wmatrix only provides frequency values for the lexical items contained in a key semantic domain (Table 9).

Word	Semtag	Frequency	Relative Frequency	
teacher	P1	304	0.20	Concordance
students	P1	215	0.14	Concordance
school	P1	187	0.12	Concordance
lesson	P1	143	0.10	Concordance
student	P1	133	0.09	Concordance
class	P1	114	0.08	Concordance
teachers	P1	94	0.06	Concordance
classroom	P1	84	0.06	Concordance
teaching	P1	75	0.05	Concordance
mentor	P1	50	0.03	Concordance

Table 9 Top 10 most frequent lexical items in P1 key semantic domain (ALE)

However, despite the fact that individual lexical items in a key semantic domain may not be keywords on their own, their frequency values assigned to a particular sense accumulate to the frequency value of the semantic domain, which is responsible for its keyness status through

corpus comparison. The analysis of the key semantic domains shall thus be based on the analysis of the lexical items in the domains since it is the general likelihood ranking (Rayson 2002) of the lexical items and the contextual rules (Gale et al. 1992) that constrain the focus to a particular sense.

The sampling procedure started with imposing a cut-off point for statistical significance of keyness. As mentioned earlier, Wmatrix measures statistical significance by calculating the log-likelihood statistic. To maximize the confidence level for the frequency difference not being made by chance, and to ensure that the analysis is not overwhelmed by too many key semantic domains retrieved, I set a strict cut-off point for statistical significance: the critical value of 15.13 which equals $p<0.0001$.

Note that statistical significance may not be the solely legitimate criterion to determine whether a keyword (or a key semantic domain) should be selected for detailed analysis. It tells how much evidence one could have for a difference between two corpora but cannot reveal how big or important a given difference is. As suggested by Andrew et al. (2011), "just because a particular test is statistically significant does not mean that the effect it measures is meaningful or important" (p. 60). One may need an indicator of the magnitude (Rosenfeld and Penrod 2011) of the statistically significant difference -- the effect size. I chose Log Ratio as a measure of calculating effect size to represent how big the difference is between two corpora (Hardie 2014) for a key semantic domain (Figure 8).

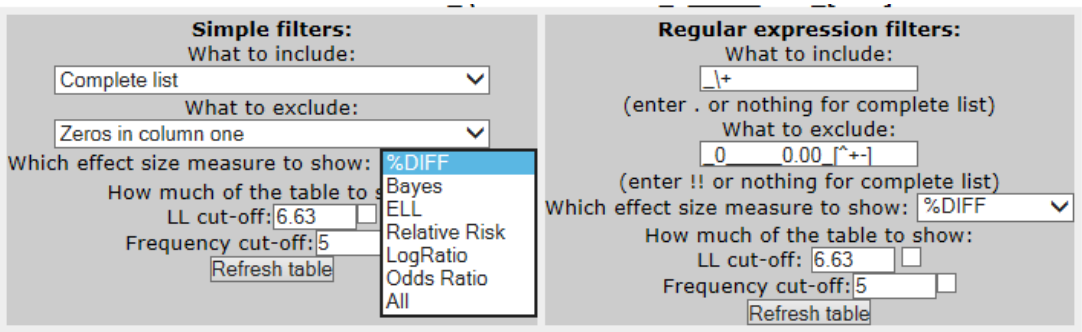


Figure 8 Available effect size measures in the keyword procedure

The basic idea is to divide the normalized frequency of a key semantic domain in the reference corpus by its normalized frequency in the focus corpus, which will give us a ratio. The log ratio of the key semantic domain is obtained by converting this ratio into a logarithm. As summarized by Hardie (2014), "every extra point of Log Ratio score represents a doubling in size of the difference between two corpora, for the keyword under consideration". A log ratio 0 means the domain has a same relative frequency in the two corpora. A log ratio larger than 0 means the domain is over-used in the focus corpus, while a log ratio smaller than 0 means the domain is under-used in the focus corpus. For example, in ALE sub corpus, the key semantic domain *P1 Education in General* has the keyness value 1643.65 (LL). Its log ratio is 1.55, which means that this domain is $2^{1.55}$ times more common in the focus corpus than in the reference corpus.

In this study, I am particularly interested in the key semantic domains which are over-used in the focus corpora, I thus selected all the key semantic domains whose log ratios are larger than 0. I copied and pasted the content of the table into an MC-Excel file so that I could sort the order the items according to their log ratios. Table 10 shows the 44 key semantic domains selected in ALE sub corpus:

Ranking	Item	O1	1%	O2	2%	Over/under-use	LL	Log Ratio	Description
1	P1	2109	1.41	16329	0.48	+	1643.65	1.55	Education in general
2	Z8	26828	17.92	499076	14.71	+	937.93	0.29	Pronouns
3	E4.1+	457	0.31	1548	0.05	+	873.24	2.74	Happy
4	Z4	9021	6.03	152983	4.51	+	659.02	0.42	Discourse Bin
5	Q2.1	2526	1.69	32756	0.97	+	625.83	0.81	Speech: Communicative
6	E6+	116	0.08	362	0.01	+	235.57	2.86	Confident
7	X2.1	1648	1.1	25570	0.75	+	199.48	0.55	Thought, belief
8	S1.1.2+	210	0.14	1552	0.05	+	175.64	1.62	Reciprocal
9	Z6	2869	1.92	50387	1.48	+	163.99	0.37	Negative
10	A13.5	270	0.18	2448	0.07	+	160.76	1.32	Degree: Compromisers
11	A5.4+	524	0.35	6342	0.19	+	160.27	0.91	Evaluation: Authentic
12	X3.2-	73	0.05	244	0.01	+	140.91	2.76	Sound: Quiet
13	A14	1278	0.85	20464	0.6	+	131.42	0.5	Exclusivizers/particularizers
14	X2.2+	1407	0.94	23019	0.68	+	128.26	0.47	Knowledgeable
15	X2.3+	179	0.12	1487	0.04	+	124.48	1.45	Learning
16	S1.1.3+	134	0.09	936	0.03	+	121.54	1.7	Participating
17	X3.2	228	0.15	2372	0.07	+	102.38	1.12	Sensory: Sound
18	X8+	348	0.23	4378	0.13	+	94.9	0.85	Trying hard
19	N4	1539	1.03	26860	0.79	+	92.15	0.38	Linear order
20	A7+	2649	1.77	49406	1.46	+	90.29	0.28	Likely
21	X4.1	639	0.43	9645	0.28	+	88.2	0.59	Mental object: Conceptual object
22	A1.1.1	2243	1.5	41892	1.23	+	75.4	0.28	General actions / making
23	I3.2+	55	0.04	340	0.01	+	58.58	1.87	Professional
24	E6-	170	0.11	2014	0.06	+	55.22	0.94	Worry
25	Q2.2	1351	0.9	24771	0.73	+	54.14	0.31	Speech acts
26	A7+++	88	0.06	800	0.02	+	52.13	1.32	Likely
27	E2+	419	0.28	6496	0.19	+	50.92	0.55	Like
28	A3+	7747	5.17	161836	4.77	+	48.1	0.12	Existing
29	A1.8+	185	0.12	2410	0.07	+	45.15	0.8	Inclusion
30	A13.6	215	0.14	2976	0.09	+	42.44	0.71	Degree: Diminishers
31	X2.2	97	0.06	1026	0.03	+	41.94	1.1	Knowledge
32	A1.2+	66	0.04	598	0.02	+	39.35	1.32	Suitable
33	A7	284	0.19	4320	0.13	+	37.82	0.58	Probability
34	I3.1	404	0.27	6676	0.2	+	34.72	0.46	Work and employment: Generally
35	Q1.3	87	0.06	1032	0.03	+	28.16	0.93	Telecommunications
36	N5-	269	0.18	4348	0.13	+	26.27	0.49	Quantities: little
37	Z7	980	0.65	18822	0.55	+	24.35	0.24	If
38	X9.1+	172	0.11	2598	0.08	+	23.66	0.59	Able/intelligent
39	A1.5.2+	45	0.03	450	0.01	+	22.03	1.18	Useful
40	B2	32	0.02	276	0.01	+	20.86	1.39	Health and disease
41	X5.2+	273	0.18	4620	0.14	+	20.22	0.42	Interested/excited/energetic
42	S6+	906	0.61	17591	0.52	+	19.73	0.22	Strong obligation or necessity
43	X9.1	50	0.03	550	0.02	+	19.7	1.04	Ability and intelligence
44	N6	99	0.07	1398	0.04	+	18.11	0.68	Frequency

Table 10 Key semantic domains (ALE)

In this study, both sequences of lexical words (e.g. *student teacher*) and sequences of lexical and function words (e.g. *[verb] students to [verb]*) are worth noticing, while those whose components are frequent than rare are of much interest to reveal language patterns. Such word-sequences also have greater potential to show features of discourse frames instead of exclusively focusing on highly technical terms. This is the reason why I chose Log Likelihood as the statistical metric to calculate collocational strength between lexical items selected within the key semantic domains. This is consistent with the metric to determine statistical

significance for the keyword procedures and explains why a threshold (minimum frequency of 5) is set for lexical item selection within the key semantic domains. A small number of lexical patterns were selected. The patterns were searched and analyzed through the concordance lines and stretched talk scenarios were then selected to be further transcribed for detailed analysis.

4.3.4 *Further data transcription*

The small group talk sessions in NUCASE were broadly transcribed. The whole small group talk sessions containing the selected scenarios were listened to through Audacity, a free audio software program. The selected scenarios were then further transcribed into Microsoft Word files following the transcription conventions established and developed by Gail Jefferson[15] (2004).

Scholars wrote extensively on the significance of transcribing processes and transcriptions to the research on human interaction (see Psathas and Anderson 1990; Jefferson 1985, 1996, 2004; Seedhouse 2004a; Duranti 2006; Liddicoat 2007; ten Have 2007; Hutchby and Wooffitt 2008; Jenks 2011). To summarize, there are a number of reasons spoken interaction needs to be transcribed into a written form for data analysis:

1. It is easier for the researcher working on written transcripts to focus on certain small, delicate parts of the data, especially the ones prone to neglect, overlooking or even distortion. As ten Have (2007) suggested, the researcher can "freeze" the data into written transcripts for repetitive analysis.
2. The transcribing process is equally relevant and important to data analysis as the transcript itself. Multiple hearings of recordings contribute to more objective observations, more careful

[15] See Appendix C for a partial list of Jeffersonian conventions

considerations and more accurate representation of what is really happening in the data (see Garfinkel 1967; Sacks 1984).

3. Details of the spoken data can be represented in written transcripts at different levels according to the researcher's interest and decision. Standardized transcribing symbols (e.g. Jeffersonian conventions, Du Bios' symbols) contribute to practical dissemination and communication of information (Brandt 2011) in/out of the academic circle by visually and conceptually symbolizing features of human interaction.

In this study, adding detailed interactional features to the broad transcripts is of particular relevance in the following aspects:

1. As the researcher in this study is not the transcriber of the broad transcripts, listening to and further transcribing the recorded spoken data helps her get more familiar with the originally used conventions and identify and correct errors in the broad transcripts.
2. The transcribing process helps the researcher to identify interesting parts of the data and make better decisions concerning how the features are to be properly represented.
3. Further transcribing adds more interactional details of small group talk which are highly relevant to contextualized analysis, e.g. prosodic features such as intonation, stress, differences in volume, pitch and length of sound (Liddicoat 2007); temporal features such as overlapping talk, pauses and gaps between turns (Brandt 2011).

For the convenience of data analysis and for the consistency in the presentation of research findings, several symbols in the broad transcripts are adapted into the Jeffersonian conventions (Table 11).

Symbols in broad transcripts	Symbols in Jeffersonian transcription	Denotation of symbols
full stop .	excluded since full stop . indicates a micro pause	a completed intonation unit
plus mark +	dash -	interruption or cut-off
equal sign =	dash -	unfinished words or single letters
<$=>… </$=>	not specified	unfinished sentences
<$H>…<$/H>	(word) parenthesized words and speakers	the transcriber's guess on dubious speaker utterance
<$G>…</$G>	empty parentheses ( )	unintelligible or inaudible utterance
<$OL>…</$OL>	left and right brackets [ ]	overlapping talk
<$E>…</$E>	Double parentheses (( ))	the transcriber's comments or extra-linguistic information

Table 11 Symbols in broad transcripts adapted to the Jeffersonian convention

Note that transcripts, however detailed, are not neutral and objective representations of talk (Green et al. 1997; Liddicoat 2007). Transcripts involve subjective researcher decisions at both cognitive and action levels thus reflect constructed rather than authentic realities of human interaction. The subjective decisions in turn influence how the researcher perceives and approaches the interactional structures by highlighting some features while diminishing others (Liddicoat 2007) in the ongoing analysis (Gumperz and Berenz 1993) for different purposes at different times (Mishler 1991).

4.4 Reliability, validity and generalizability

4.4.1 *Reliability*

Reliability refers to the repeatability and replicability of analytic findings in similar research contexts, or in other words, whether the methods and procedures applied in a study can give

the same result on successive trials. Both quantitative and qualitative research should be presented in a way that could bring a level of transparency to the analytic claims made (Nikander 2008), which enables other researchers "to make their own checks and judgments" (Potter and Edwards 2001: 108). In this study, the quantitative analysis conducted using CL techniques ensure the reliability of the analytic claims since the frequency calculations and keyword procedures are strictly based on standardized statistical tests and the results are represented in a consistent manner. However, the researcher should be aware that slight differences may exist when different corpus software is used and that striking differences might appear if different statistical metrics are chosen. Other researchers may choose different computer software to approach the same data set due to their accessibility and may use different statistical metrics according to their specific research focuses.

In addition, how the linguistic features generated using CL approaches would be selected to provide a manageable sample for subsequent analysis can be highly arbitrary and optional among different researchers. In terms of the more qualitative analyses through the concordance examinations, the issue of reliability can be more complicated. What is central to the complication is the quality of the transcripts of the audio- and video-recorded data in the corpus. A transcript should provide a detailed and accessible representation of the interaction under investigation. However, no matter how the researcher would try to make the transcript as detailed as possible, it is never able to reflect the whole picture of the interaction without losing a single communicative feature. Due to the unavailability of video recordings for most of the small group talk sessions and the relatively poor quality of the available ones, the reliability of the transcripts in NUCASE has been significantly reduced since a large number of non-verbal modalities are untraceable from the transcripts. To add to the transparency of the analytic claims and encourage more critical scrutiny, the researcher can provide the

original recordings along with the transcripts in the result presentation, which would help to increase the reliability of the research findings.

4.4.2 *Validity*

In his work on qualitative research featuring micro analyses on interaction, Seedhouse (2004b) discusses three types of validity: internal, ecological and construct. Internal validity refers to the credibility of research findings. As mentioned in the previous section, the increasing availability of data on language use in context and the improvement of transcribing accuracy allow the research community to check the credibility of the analytic claims made in individual studies.

Construct validity, when examined within an emic methodological framework, refers to the ways in which participants organize language use and social relations in interactions, emphasizing the appropriateness of inferences of contextualized intentions based on observations, meaning negotiations and sequential designs. As argued in the previous chapter, the status of context as an interface to combine Corpus Linguistics and Interactional Linguistics analyses corresponds well to the relationship between language use and social identities, thus ensures the construct validity of the research findings. In addition, the methodological synergy applied in this study can improve the validity of the research outcomes through "the convergence and corroboration of the findings" (Dörnyei 2007: 45).

4.4.3 *Generalizability*

Generalizability refers to the extent to which analytic findings can be applied to other research settings beyond the one where they are generated. As is widely acknowledged, quantitative

research is characterized by a quest for generalizability where "numbers, variables, standardized procedures, statistics and scientific reasoning" are all part of such a quest "for facts that are generalizable beyond the particular and add up to wide-ranging, ideally universal laws" (Dörnyei 2007: 34). On the other hand, qualitative research has been criticized for a lack of generalizability due to its exploratory nature, emergent research design, insider perspectives, samples of small sizes and interpretative analyses (Dörnyei 2007). Even though some cases, when examined using qualitative methods, can provide insights into certain phenomenon thus can be taken as representative, there is a risk of overestimation (or as Yates [2003]: 224 calls the "potential over-reading") while the specific conditions may not apply to others in a broad domain (Duff 2006). However, it is only partially true that the potential of qualitative findings to be generalized is highly limited. The social and institutional perspective adopted in this study reveals details of the participants' contextual orientations and broad expectations, at least within similar institutional settings, it thus can to some extent generalize the features of language use at the discursive level.

4.5 Summary

This chapter addressed issues of the research design of this study. Brief descriptions of the main data set (retrieved from NUCASE), the sub corpora and the reference corpus were provided. Procedures of data analysis were articulated. The validity, reliability, and generalizability of research results were intensively discussed, with ethical issues taken into account. The sampling procedure applied in this study to select target key semantic domains and lexical items for data analysis is summarized as follows:

1. Set the measure of determining statistical significance: Log Likelihood;
2. Impose the cut-off point for statistical significance: critical value 15.13 (p<0.0001);

3. Set the measure of effect size: Log Ratio;
4. Run keyword procedures for the sub corpora (RC: MICASE);
5. Select key semantic domains that are over-used in the focus corpora (Log Ratio>0);
6. Exclude any irrelevant domains based on the research focus;
7. Select lexical items with a minimum frequency of 5;
8. Select target lexical concepts for further analysis.

The sampling procedure, as the very first stage of data analysis, shows that

- The keyword lists are created in enough detail so that other researchers can replicate the procedure;
- The sample is selected based on the research focus and the nature of data indicated by statistical measures, thus well justified with a clear rationale;
- The sample selection process involves subjective decisions of the researcher thus cannot offer a free pass to objectivity (Baker 2015);
- Computers and automated tagging systems are not error-proof thus the researcher's reflexive intuition is sometimes necessary.

This study follows the analytic sequences featuring a linear order from Corpus Linguistics (CL) analysis to Interactional Linguistics (IL) analysis of the university small group talk in terms of knowledge construction (Figure 9).

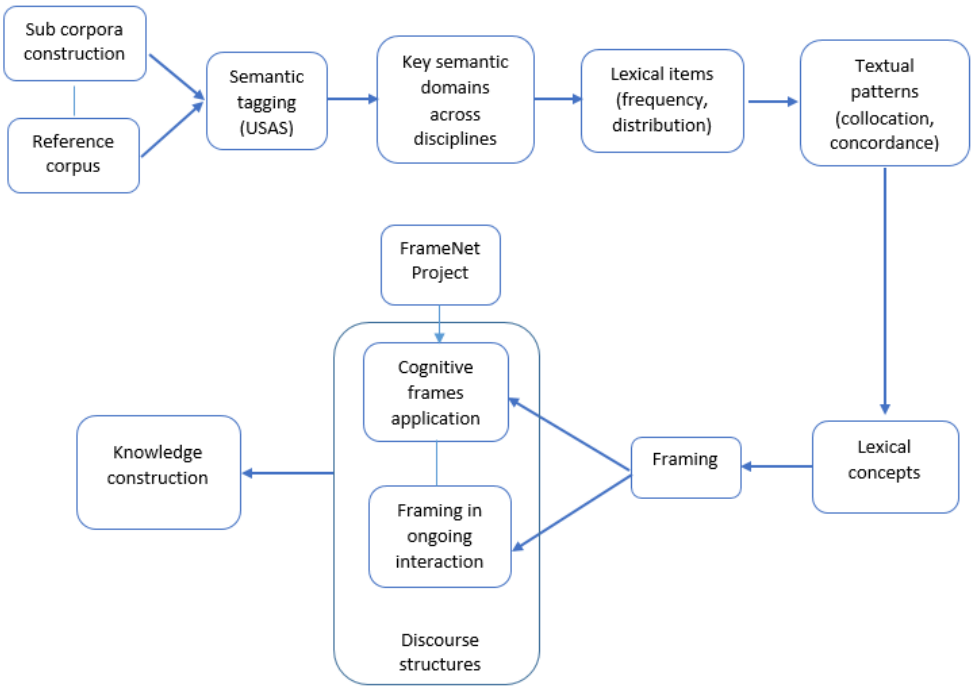


Figure 9 Analytic sequences of NUCASE inquiry

The analytic sequences are adopted because:

- Empirically, consideration of the quantitative aspects of the participants' lexis use in naturally occurring interaction can provide the researcher with a macroscopic view of the language phenomena at the semantic level. Attending to the qualitative aspects of how the lexical concepts are represented and negotiated within micro-discourse structures contribute to a minute description of interactional features for framing. In addition, starting the inquiry with a manual sorting and identification of any potentially analyzable units of meaning can be extremely time-consuming and would risk subjectivity in initial data analysis.
- Methodologically, the frequency-based inquires of lexis use involving statistical significance tests (keyword procedures, collocations) add to the reliability of the research design, thus can ensure the repeatability and replicability of findings in similar research

contexts. Key semantic domains generated from the sub corpora represent the key concepts across the disciplines. Lexical items are then selected for further analysis, with their associative lexical concepts revealed from the concordance lines and collocational patterns. These lexical concepts are taken as discursive signposts to provide a way-in to the local contexts where framing behaviors for knowledge construction can be observed in detail.

Chapter 5. Research Findings

5.1 Introduction

This chapter reported the research findings of how the participants frame their small group talk to collaboratively construct knowledge. It is argued that motivating contexts (Fillmore 1982a) play an important role in shaping patterns of framing through participants' talk of lexical concepts. The contextual resources (Linell 1998a) managed by the participants contribute to situated meaning construction. These include immediate contextual resources such as the prior discourse in the group encounter (co-text) and mediate contextual resources such as the participants' shared disciplinary knowledge, nature and procedure of the group task, communicative routines and action types, role assignment and labour division, etc. The combinational configurations of relevant contextual resources determine what semantic potentials of lexical concepts are to be selected as access sites to particular conceptual knowledge structures (Evans 2006).

Section 5.2 gave an overview of the key semantic domains in each sub corpus and presents examples to illustrate how lexical concepts encoded by words provide access to coherent, non-linguistic knowledge structures (evoking frames). Section 5.3 focused on how the participants invoked interpretive frames to assign the target discourse relevance and coherence. Collocational patterns and concordance lines of distinct lexical concepts associated with particular words showed evidence of situated meaning construction. Corpus Linguistics approaches were applied in these two sections, with both quantitative and qualitative analyses conducted. Section 5.4 demonstrated how expanded scenarios of small group talk can reveal patterns of framing-in-interaction for knowledge construction. The relationships between discourse, speaker roles, motivating contexts and the process of knowledge construction were discussed. Interactional Linguistics analysis was conducted in this section to examine the

sequential organization of the participants' talk, with a broad range of contextual resources taken into account.

5.2 Words evoke frames: lexical concepts across the key semantic domains

Based on the comparison of the semantically tagged focus corpus (NUCASE) and the reference corpus (MICASE), the key semantic domains in each of the three sub corpora were identified. These semantic domains revealed what aspects of meaning are shared by the lexical concepts encoded by specific linguistic forms. As discussed in Section 4.3.3, only the over-used, non-shared key semantic domains were selected for further analysis. They are the most salient discursive signposts to reveal what the talk in each sub corpus is mainly about and how the epistemic domains are distinct from each other. Table 12 shows the non-shared key semantic domains in the sub corpora, with cases due to mis-tagging excluded.

Sub corpus	**Non-shared key semantic domains (LL)**	**Example of cases due to mistagging (excluded)**
ALE	*P1 Education in general (1643.65).*	N/A
BI	*Y2 Information and technology and computing (1566.77);* *B3 Medicines and medical treatment (80.87).*	e.g. The word *flow* is mistagged into *M4 Sailing and swimming*. *work* and *flow* show a strong collocational association in the sub corpus, thus *flow* is more appropriately to be tagged into *Q1.2 Paper documents and writing*, or *Y2 Information technology and computing*. *M4* is thus excluded.
ME	*M4 Sailing and swimming (2393.58);* *W4 Weather (1243.42);* *O3 Electricity and electrical equipment (1185.10);* *W3 Geographical terms (748.00);* *C1 Arts and crafts (229.28);* *O4.6 Temperature (+153.86/-92.64).*	e.g. The word *lit* is mistagged *into O4.6+ Temperature: hot/on fire*. However, it is used in this sub corpus as a shortened form of *literature*, since it shows a strong collocational association with the word *review* (all cases as shown in concordance lines). The word *lit* is thus excluded.

Table 12 Non-shared key semantic domains across the sub corpora

5.2.1 *Applied Linguistics and Education (ALE)*

Applied Linguistics and Education is an interdisciplinary field which is aimed to identify, investigate and offer solutions to language-related educational practices. As shown in Table 10, the key semantic sub domain *P1 Education in General* has the largest keyness value (LL: 1643.65) and is exclusively used in ALE sub corpus. *P1* can be seen as a body of lexical concepts associated with words which are clustered based on the approximation of their semantic values. The coherent body of knowledge is derived from the real-word experience of teaching and learning. The internalized mental representation (Clausner and Croft 1999) then invokes schematizations (Petruck 1996) concerning how categories become relevant and subject to represent stereotyped situations (Minsky 1974) in such practice. The categories, taking the form of linguistic forms, are expected to be assembled through conventional language repertoire but in non-conventional patterns to reveal the unique meaning of certain utterance (Evans 2006). In this section, one shall see how the lexical concepts involved in ALE sub corpus contributed to revealing discursive features at both the conceptual and action levels.

Word	Semtag	Frequency	Relative Frequency
teacher	P1	304	0.20
students	P1	215	0.14
school	P1	187	0.12
lesson	P1	143	0.10
student	P1	133	0.09
class	P1	114	0.08
teachers	P1	94	0.06
classroom	P1	84	0.06
teaching	P1	75	0.05
mentor	P1	50	0.03
lessons	P1	48	0.03
tutor	P1	41	0.03
teach	P1	39	0.03
schools	P1	36	0.02
pupils	P1	34	0.02
learners	P1	27	0.02
classes	P1	27	0.02
student_teacher	P1	21	0.01
thesis	P1	20	0.01
learner	P1	18	0.01
training	P1	17	0.01
education	P1	17	0.01
classrooms	P1	17	0.01
head_teacher	P1	16	0.01
taught	P1	15	0.01
geography	P1	15	0.01
tutors	P1	11	0.01
exam	P1	10	0.01
module	P1	10	0.01
phd	P1	9	0.01
year_nine	P1	9	0.01
educational	P1	8	0.01
academic	P1	8	0.01
university	P1	8	0.01
pedagogical	P1	8	0.01
mentors	P1	7	0.00
geography_department	P1	7	0.00
pgce	P1	7	0.00
exams	P1	6	0.00
test	P1	6	0.00
seminar	P1	6	0.00
pupil	P1	6	0.00
year_eight	P1	6	0.00
study	P1	6	0.00
train	P1	5	0.00
testing	P1	5	0.00
assignment	P1	5	0.00
student_teachers	P1	5	0.00
year_ten	P1	5	0.00
year_seven	P1	5	0.00
homework	P1	5	0.00

Table 13 Lexical items (freq. >/= 5) in P1 (ALE)

Table 13 shows the 44 linguistic forms included in *P1* whose frequency is no less than 5. Most of these linguistic forms are used to represent institutional actors (e.g. teacher, student, pupil, tutor, learner, etc.) and activities (e.g. teaching, training, seminar, test, assignment, etc.). Lexical concepts encoded by the linguistic forms are believed (by the tagging system) to provide access to the knowledge of "education in general". In *P1* the behaviours of *teaching and learning* are explicit, which creates the motivating context to evoke the [EDUCATION]

frame with the knowledge about education locally constructed in the participants' talk. Table 14 shows the elements of the [EDUCATION] frame.

	Teacher	Learner	Content	Location/Means
Teach	Subject	(to; D-object)	[D-object]	[PP]
Tutor	Subject	D-object	[on]	[PP]
Train	Subject	D-object	[on]	[PP]
Test	Subject	D-object	[on]	[PP]
Mentor	Subject	D-object	[on]	[PP]
Learn	(from)	Subject	D-object	[PP]

Table 14 The semantic and syntactic valence[16] (active voice) of the [EDUCATION] frame in P1 (structure adapted from Fillmore and Atkins 1992: 79)

The word *subject* in a cell indicates that in the potential use of the verb represented by its row, the verb's subject represents the category labelled by its column. D-object means direct object. I-object means indirect object. A preposition indicates a prepositional phrase it heads to represent the element. PP means prepositional phrase. Parentheses mark optionality and square brackets mark contextualized omissibility (Fillmore and Atkins 1992). Table 12 shows a preference for the teacher's domination in educational institutions while the learner's motivated involvement is to a large extent missing in the frame. The imbalance of the inner-frame prominence may be partially due to the word sense disambiguation (WSD) process applied in the USAS tagging system. In the tag set most semantic senses related to learning mentality or behaviours are tagged within *Domain X Psychological action, states & processes*, e.g. *X 2.2 Knowledge; X 2.3 Learn; X 2.4 Investigate, examine, test, search; X 2.5 Understand;*

[16] "Valence" refers to the particular kinds of constituents, in terms of semantic roles, grammatical functions, and phrase types, with which a word combines in a grammatical sentence (FrameNet Glossary).

X 4.2 Mental object: means, method, etc. While it is unclear whether the tagging system treats "learning" more as a psychological/cognitive concept and "teaching" as a procedural one (see Schraw and Olafson's [2003] suggestion that teaching is more process-oriented than content-oriented compared with learning), it does highlight the importance of inter-frame examinations to understand how distinct the semantic frames evoked by interrelated lexical concepts can be and what interpretive consequences the distinction may have.

Patterns of meaning construction for knowledge representation can also be revealed from words' collocational behaviours. Table 15 shows the most salient collocational patterns (collocates hitting no less than 5 times within the ±5 spans of the search-word) of the selected word sequences in *P1* which constitute only lexical items (Biber 2009).

N	Word	With	Relation	Texts	Total	Total Left	Total Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	PLAN	lesson	48.35	4	13	2	11			1	1			10			1	
2	HEAD	teacher	39.20	4	17	16	1					16						1
3	FIRST	lesson	30.25	3	7	6	1			1		5					1	
4	NEXT	lesson	30.24	3	9	6	3					6				1	2	
5	YEAR	class	28.85	3	16	15	1	2			12	1						1
6	WHOLE	school	25.55	3	14	13	1	2	1	1	1	8						1
7	PROFESSIONAL	tutor	24.57	2	6	6	0					6						
8	OTHER	students	21.14	3	11	11	0	1	1		1	8						
9	*EACH	lesson	19.90	2	8	8	0					8						
10	STUDENT	teachers	19.29	2	7	6	1	1				5						1
11	TEACHER	student	12.28	4	24	1	23	1						22		1		
12	STUDENT	teacher	12.28	4	24	23	1			1		22						1

Table 15 Lexical collocates (LL) of selected linguistic forms from P1 (ALE)

Noun phrases are found to have great potential to show collocational patterns of associated lexical concepts. In fact, there are far more nouns used in *P1* than verbs and adjectives. Nouns "contain information about several domains and thus represent categories" (Gärdenfors 2014: 116-117). Most of the nouns in *P1* can be categorized as *relational nouns*: e.g. *teacher, student, learner, classroom, seminar, university, exam, assignment*, etc. For such a noun, its membership of a semantic category has to be determined by verifying its relations to other

entities, since locating its intrinsic properties does not suffice to identify it (Asmuth and Gentner 2005). For example, *teacher* cannot be sufficiently identified or categorized if it is not understood against the lexical concept associated with *learner*. The relational category can be characterized in terms of a common relational structure to determine the membership, thus is "linguistically constructed" (Gentner and Kurtz 2005: 156).

Sharing some commonalities with verbs and prepositions, relational nouns show greater sensitivity to context (Asmuth and Gentner 2005; see also Kersten and Earles 2004). The contextual sensitivity of relational categories reveals the relationship between the property of a lexical concept and the context it is embedded in. This is perceived as the "context-dependent property" (Barsalou 1982: 83) of relational categories. For example, in the [EDUCATION] frame the central meaning element carrying the defining feature(s) of the frame can be articulate and explicit (*teaching and learning*), while the paths of access to the central meaning element provided by the lexical concepts encoded by some linguistic forms can vary in local contexts. For instance, *university* encodes the mental representation of a physical location or a conceptual institution; *assignment* encodes the mental representation of a pre-described obligation/expectation; and *exam* encodes the mental representation of a procedure/means, etc. Different knowledge structures or different facets of a knowledge structure can be activated thanks to the semantic potential of lexical concepts (Evans 2006) which serves situated meaning construction. In such sense, the "segment of reality" (Brinton 2000: 112) a frame as a conceptual entity denotes is not only the *conceptual* reality (schematized human experience) but also the *contextual* reality (paths of access to schematization). The context-dependent property of lexical concepts reveals patterns of meaning construction by semantically encoding contexts rather than through respective words (Barsalou 1982).

Selection of lexical concepts can also reveal the structure of action in certain discourse. This is reflected in the *Micro Analysis Research Group* (MARG)[17] talk in ALE sub corpus. Inspired epistemologically and methodologically by Conversation Analysis (CA), the MARG talk is framed procedurally by an intensive use of CA-related terms, the explicit conceptual reflection and the identification of analytical focuses through sequential observations of spoken data. Table 16 shows the most salient collocational patterns of CA-related terms in the MARG talk (collocates hitting no less than 5 times within the ± 5 spans of the search-word). Word sequences are listed in descending order according to the strength of collocational association.

[17] MARG is a research team constituting a group of (mainly research) students and academic staff who meet on a weekly basis to jointly analyze video and audio data of real life spoken interaction. At a MARG meeting, participants discuss issues of social/institutional interaction based on their analysis of the original and transcribed video and audio data drawing theories and methods of Conversation Analysis (CA).

N	Word	With	Relation	Texts	Total	Total Left	Total Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	SEVENTEEN	line	24.06	3	3	0	3							2	1			
2	IRF	irf	22.51	2	16	1	1		1				14				1	
3	DELAYED	pair	21.30	2	6	5	1				5					1		
4	FOUR	line	18.43	4	10	1	9				1			2	3	3		1
5	SEVEN	line	18.27	4	22	2	20		1	1				6	10	2		2
6	NINE	line	18.27	4	20	3	17	1	1		1			5	10	1		1
7	FIVE	line	17.54	4	15	0	15							5	9	1		
8	THIRTEEN	line	17.54	2	3	0	3							2				1
9	FORTY	line	17.51	3	7	0	7							5			2	
10	ADJACENCY	pair	16.89	2	14	14	0					14						
11	PARTS	pair	16.63	2	14	2	12		1	1				11				1
12	PAIR	parts	16.63	2	14	12	2	1				11				1	1	
13	EIGHTEEN	line	16.36	2	3	0	3							3				
14	FOR	seconds	15.93	4	13	13	0		6	1	6							
15	THIRTY	line	15.70	4	31	3	28	1		2				20	2	1	4	1
16	TWENTY	line	14.51	4	33	3	30	1	1			1		25	1	2	1	1
17	HALF	page	14.18	2	19	17	2	2			1	14				2		
18	LINE	three	13.65	4	15	15	0	3	1	3	6	2						
19	CHANGE	shift	13.05	2	16	4	12			2	2				6	4		2
20	SECOND	pair	10.75	2	21	20	1	1			2	17					1	
21	PART	pair	10.48	2	20	2	18					2		18				
22	TWO	line	10.01	4	18	2	16					2		3	8	2	2	1
23	FIRST	page	9.83	2	7	7	0		1			6						
24	SEVENTY	line	9.27	2	6	0	6							6				
25	FIRST	pair	9.20	2	12	12	0					12						
26	NEW	topic	9.19	2	6	5	1					5				1		
27	SIXTY	line	8.82	2	8	0	8							8				
28	STEPWISE	topical	8.71	1	10	10	0	1		1		8						
29	TOPICAL	stepwise	8.71	1	10	0	10							8		1		1
30	STEPWISE	movement	8.61	1	9	9	0		1		8							
31	MOVEMENT	stepwise	8.61	1	9	0	9								8		1	
32	TOPICAL	movement	8.15	1	22	19	3	1				18				1	1	1
33	MOVEMENT	topical	8.15	1	22	3	19	1	1	1				18				1
34	MOVEMENT	boundary	8.05	1	14	4	10	1		2	1				6	3	1	
35	BEFORE	question	6.90	2	11	0	11							10		1		
36	OF	sequence	6.60	3	10	10	0		1	2	7							
37	TOPIC	topic	5.62	3	254	36	36	5	10	15	6		182		6	15	10	5
38	AT	line	4.89	4	39	36	3	2	3	1	1	29			1	2		
39	THE	topic	4.79	2	120	79	41	12	13	12	7	35		6	11	14	4	6
40	AND	shift	4.45	3	18	7	11		2	2	2	1		7		1	2	1
41	MANAGEMENT	topic	3.99	1	22	3	19	1	1		1			14	2		2	1
42	ON	line	3.79	4	28	23	5	1	1	8	3	10		1			2	2
43	IN	transcript	3.25	2	7	7	0			1	6							
44	VERTICAL	line	3.01	1	26	26	0					26						
45	LINE	vertical	3.01	1	26	0	26							26				
46	TOPIC	shift	2.88	1	38	23	15	1	1	2	2	17		3	8		1	3
47	SHIFT	topic	2.88	1	38	15	23	3	1		8	3		17	2	2	1	1
48	THE	transcript	2.17	3	26	23	3		1			22			1	2		
49	CHANGE	topic	2.00	1	37	13	24		2		3	8		15		5	3	1
50	THE	line	1.97	4	52	39	13	7	7	11	10	4		1	2	4	4	2
51	THIS	line	1.94	4	17	10	7	1	5	3		1			4	1		2
52	TO	topic	1.28	2	39	32	7	10	2	9	7	4		2	1	2	2	
53	IS	topic	1.23	1	58	30	28	2	11	4	5	8		8	9	5	3	3
54	FIRST	question	1.22	1	6	6	0					6						
55	A	page	1.20	2	24	18	6		1		15	2			1		3	2
56	ABOUT	topic	0.82	2	19	13	6	1		2	3	7		1			2	3
57	IN	line	0.61	4	53	44	9	4	3		2	35		3	2	1	1	2
58	OF	topic	0.57	2	40	31	9	4	3	7	8	9		2	1	1	1	4
59	A	line	0.37	3	51	45	6	3	5	2	30	5		1	1	2	1	1
60	A	vertical	0.22	1	28	28	0	1				27						
61	THAT	line	0.19	4	37	20	17	4	3	4	5	4		4	2	6	2	3
62	A	topic	0.18	3	47	31	16	6		3	8	14			7	5	1	3
63	THE	boundary	0.06	1	14	11	3	2		1	2	6		2	1			
64	THE	pair	0.04	2	24	21	3	2	2	1	15	1				2		1
65	ON	topic	0.02	1	23	20	3	4	2	2	5	7			1			2

Table 16 CA-related terms in MARG talk (P1_ALE)

65 collocational structures are identified concerning the explicit use of CA-related terms which can be roughly categorized into the following patterns:

- Textual reference: e.g. *line*+no.; *the/this/a/that line*; *half/first page*; *in/the transcript*, etc.
- Structural identification: e.g. *movement/the boundary; IRF; adjacency/first/second/the pair; pair parts; sequence of; first question; question before,* etc.

- Analytical focus: e.g. *topical movement; stepwise movement; topic change/shift; change topic; delayed…pair; stepwise changes/topic,* etc.

The three patterns correspond to the action process in which the MARG participants convey and interpret meanings for knowledge construction to perceive the social reality revealed by the video/audio-recorded interaction as the target of analysis (Figure 10).

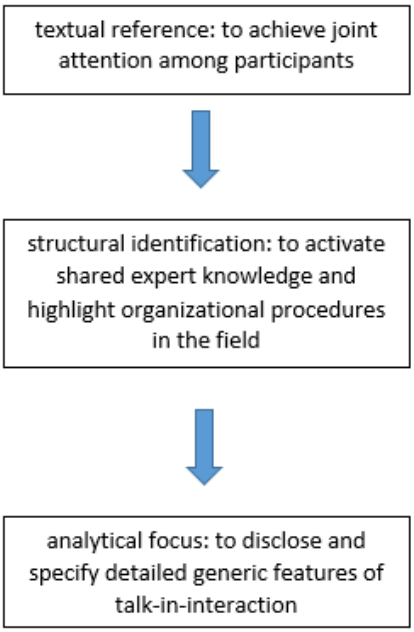


Figure 10 Action process in MARG talk

Knowledge construction in the MARG talk, guided by the participants' orientation to understanding the orderliness of social interaction, is continuously seeking relevance of meaning-making moves to the "textual reality" in the data. The observation has a consequence for interpretation of generic discourse. It allows the interpreter to assign coherence to the discourse by invoking particular frames such as [INFORMATION_DISPLAY], [RECOGANIZING AND INDICATING], [SCRTINY] from the patterned language use to understand its procedures and routines at the institutional level.

5.2.2 *Bio-informatics (BI)*

Bioinformatics is an interdisciplinary field which develops methods and software tools to understand biological data. Computer programming is part of the methodology of biological studies belonging to the field. The identification of the two sub key semantic domains in BI sub corpus is consistent with the disciplinary description:

- *Y2 Information technology and computing (Y Science & technology)*; and
- *B3 Medicines and medical treatment (B The body & the individual).*

The two sub key semantic domains, however, are not ranked next to each other in terms of keyness value. *Y2* hits the top of the list with the keyness value of 1566.77 (LL), while *B3* is ranked the 20th with the keyness value of 80.87 (LL). If one takes the view that Bioinformatics is about making sense of the biological data from a computational perspective, one may be able to tentatively explain the keyness difference by assuming that the application of computer programming as the methodology for biological data analysis is much more emphasized than the nature of the data. The statistically defined *topicality* (see Strawson 1964, Gundel 1985, van Kuppevelt 1995 for a general discussion; see Slimani 1989, Stokoe 2000, Dooley 2007 for different views), or what is called "aboutness" from the Corpus Linguistics perspective, is revealed by the discursive status of *Y2* as "more key" than that of *B3*. In this section, one shall see how lexical concepts providing access to knowledge of science and technology were locally associated with biology and medicine for situated interpretation of the interdisciplinary discourse.

A closer look at *B3* reveals "what is of current interest or concern" (Strawson 1964: 104) in the participants' talk concerning the discursive relevance of *B3* to what is acknowledged as biological attributes. A simple searching in BI sub corpus for *biology* identifies two potential patterns in which its associative lexical concepts provide access sites to specific knowledge structures:

- Pattern 1: *Y1 Science and technology in general* + *L1 Life and living things*;
- Pattern 2: *Y1 Science and technology in general* + *B1 Anatomy and physiol*ogy

Y1 belongs to the semantic domain *Y Science & technology*. *L1* belongs to *L Life & living things* and *B1* belongs to *B The body & the individual*. Biology, in both patterns, is acknowledged as a field of science. Pattern 1 is about the denotation of biology, emphasizing the lexical concept of OBJECT AND TARGET[18], while Pattern 2 is about its connotation, emphasizing the lexical concept of METHODS OF INQUIRY. Then how is the conceptual connection built between the [MEDINES AND MEDICAL TREATMENT] frame and lexical concepts associated with biology? The conceptual structures for construing biology provide clues of the relevance. Medicines and medical treatment can be perceived as biologically relevant since results of anatomical and physiological research provide theoretical and practical supports and living things are direct objects of experimentation and application.

While the relevance at this level is to a large extent taken as common sense, what is worth noticing is that the [MEDINCES AND MEDICAL TREATMENT] frame is attributively constructed via linguistically encoded mental representations. This is discursively revealed in

[18] In this study, capital letters are used to represent lexical concepts.

its close relationship with HEALTH and DISEASE (*B2 Health and disease*); in other words, the HEALTH and DISEASE as lexical concepts show a great potential to activate knowledge of medicines and medical treatment. Medicines and medical treatment are conceptualized based on human cognition, involving a series of activities such as comparing, identifying, differentiating, evaluating, classifying and affecting. The word *disease* is associated with a strongly negative evaluation as culturally constructed, with the concept of NORMALITY as the main criterion of judgment while the efficacy of medicines and medical treatment as the counter power is accordingly perceived as semantically positive. The word *medicine* thus belongs to the relational category rather than the entity one, aimed at CURING ABNORMALITY.

In *B3*, the lemma *drug* has the highest frequency value (90 hits in the singular form and 45 hits in the plural form). Table 17 shows the top 10 collocates of *drug(s)* in BI sub corpus (ranked by collocational strength calculated by LL; span +/-5). Two patterns are identified to be with strong collocational strength and frequency value: *drug targets* (LL=97.002 at the R1 position, 20 hits in total) and *drug repurposing* (LL=95.565 at the R1 position, 32 hits in total).

N	Word	With	Relation	Set	Texts	Total	Total Left	Total Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	DRUGS	drugs	129.44		11	47	1	1			1			45			1		
2	DRUG	drug	118.34		13	102	6	6		4	1	1		90		1	1	4	
3	ONE	drug	103.19		7	10	5	5			4	1				4			1
4	TARGETS	drug	97.00		11	20	0	20							19		1		
5	REPURPOSING	drug	95.57		9	32	2	30			2				28				2
6	IS	drug	58.83		4	6	3	3			1	1	1			2			1
7	TARGETS	drugs	53.65		6	11	6	5	1	2	2	1				1	1	2	1
8	ONE	drugs	48.35		3	3	1	2			1				1	1			
9	EXISTING	drugs	46.02		4	5	5	0			1		4						
10	LIST	drug	38.09		6	13	4	9	2	1		1			3	4	1		1

Table 17 Top 10 collocates of *drug(s)* in (BI)

The word *target* in this context is tagged into the semantic domain *X7+ Wanting, planning, choosing* in the USAS Semantic Tagset. The semantic domain *X* concerns psychological actions, states, and processes, while the concept of CHOOSING denoted by drug targeting is framed more procedurally in the participants' talk on the basis of the shared knowledge of the inquiry methodology specific to the discipline. In addition, the theoretically supported and practically trialed action processes, in turn, shape the talk into a unique discursive structure within which concepts and methods are explicated more directly, accurately and concisely.

The word *repurpose*, by contrast, is not accurately identified in the USAS tagging system. I then considered approaching its associative lexical concepts by examining the linguistic representation of *purpose* due to the morphological and semantic associations between the two words. The USAS tagging system provides three potential semantic tags for the word *purpose*:

- *X7 Wanting, planning, choosing (X Psychological actions, states & processes)*;
- *A2.2 Affect: Cause/connected (A2 Affect* in *A General & abstract terms)*; and
- *A1.5.1 Using (A1.5 Use* in *A1 General* in *A General & abstract terms*).

Generally speaking, the word *purpose* is expected to evoke the [INTENTION FOR ACCOMPLISHEMENT] frame, while suggests a more settled determination than *intention* and emphasizes psychological formulation over the effect of action as suggested by *aim* and *end*. The deliberateness of the psychological formulation highlights two dimensions to linguistically represent the word *purpose*: understanding of the connectedness between factors involved (logic at the abstract level) and application of proper methods to address the

connectedness so as to reach the goal. Therefore, when approaching what *repurpose* possibly means one shall be aware 1) that the prefix *re-* indicates repetition or restoration denoting the concept of doing something AGAIN, and 2) that the repetition or restoration is to make a CHANGE that would take into form something DIFFERENT or NEW. 1) aims at 2) and 2) rationalizes 1). Concerning how to make such a change, the lexical concepts encoded by *purpose* are perceptually transferable to understand *repurpose*: when we consider repurposing something, we think about figuring out a different pattern of associations between relevant factors; or, renewing the existing methods for innovation.

Here is another example. *Vaccine* is a word which is frequently mentioned in *B3*. While *medicine* and *vaccine* share the associative lexical concept of PROTECTION AGAINST DISEASE, differences between the two words cannot be simply explained by examining how related they are to each other but have to be understood within the semantic domain they are embedded in. It might be easy for one to introduce the dimension of time as a parameter for comparison: a medicine is administered after a person has got a disease and vaccination happens before a person could possibly get a disease. While this may make sense in individual cases, the *Time* dimension loses its relevance when medication and vaccination are practiced in the public domain: a medicine can be discovered or developed to address expected health risk while a vaccine is produced based on a relatively large data sample of a certain disease which has already been diagnosed. That vaccines are believed to be most effective in preventing infectious diseases reveals a notable feature of collectiveness in the developmental principle and working mechanism of vaccination. An alternative way to perceive the difference involves associating the TARGET and MEANS of the effects of medication and vaccination: a medicine is administrated to an unhealthy body so as to protect

it against disease through [TREATMENT]; a vaccine is administrated to a healthy body so as to protect it against disease through [PRECAUTION].

5.2.3 *Marine Engineering (ME)*

Marine Engineering concerns applying engineering sciences to the development, design, operation and maintenance of watercraft propulsion, on-board systems and oceanographic technology. In ME sub corpus, 6 sub-sub key semantic domains are identified under 4 semantic domains (or sub domains) (Table 18). In this section, one shall see how lexical concepts can be attributively transformed to evoke different frames to construct conceptual knowledge for the disciplinary pragmatics.

Sub(-sub) key semantic domains	Keyness value (LL)	(Sub) semantic domains in USAS
M4 Movement/transportation: water	2393.58	*M Movement, location, travel & transport*
W4 Weather	1243.42	*W The world & our environment*
W3 Geographical terms	748.00	*W The world & our environment*
O3 Electricity and electrical equipment	1185.10	*O Substance, materials, objects & equipment*
O4.6 Temperature +/-	+ 153.86; - 92.64	*O4 Physical attributes in O*
C1 Arts and crafts	229.28	*C Arts and crafts*

Table 18 sub/sub-sub key semantic domains (ME)

M4 and *O3* have the largest keyness values, which means that the lexical concepts encoded by the words in them show what the talk is mainly about. Specifically, the talk is found to be centered on the design (*C1*) of electrical system/equipment (*O3*) for watercraft propulsion (*M4*) with environmental factors (*O3, O4.6, W3* and *W4*) taken into account. The lexical

concepts involved in the key semantic domains are integrated and developed into a more intricate disciplinary structure bound by the relational ideas, which manifests features of what is called *integrated knowledge codes* (Bernstein 1971) in Higher Education.

According to Becher's (1987) theory of *knowledge domains*, Marine Engineering is identified as an "applied-hard" discipline. It means that the discipline is to a large extent purposive and pragmatic, concerned with mastery of physical environment and orients to result in products or techniques (Becher 1987). Integrated conceptual patterns are thus constructed to serve the pragmatic goals, while the basic questions in identifying the recurring structures of knowledge representation concern the task of categorization (Rey 1983; Medin and Smith 1984).

A rather simplified classification tends to categorize lexical concepts according to the methods of their initial establishment (Komatsu 1992). For example, Putnam (1975a, 1975b) coined a certain class of common nouns as *natural kind terms*[19] for they label naturally occurring things like *water* and biological kinds. Schwartz (1978, 1979) identified *nominal terms* (e.g. *bachelor*) which are used attributively (Donnellan 1977). Putnam's categorization is based on empirical discoveries while Schwartz's nominal concepts are established by social conventions (Komatsu 1992).

[19] Putnam (1975a, 1975b) argued that for *natural kind terms*, concepts in fact do not mediate the link between words and their referents. He proposed what is called a "division of linguistic labor" (Putnam 1975b: 227) that it is left to the experts to figure out what, for example, *wind*, actually is and that lay people who use the term *wind* typically have some beliefs about the kind but their concept merely reflects to some extent the experts' beliefs about the kind. This analysis leads to an important implication that well-developed concepts will more be like theories than simple descriptions.

In such a categorization, the lexical items in *W3* and *W4* can be safely classified into natural kind terms, while the ones in *C1* into nominal terms. However, the sub key domains *O* and *M* seem quite problematic to fit in, since the words calculated in the domains label affects upon and relationships between objects/events other than objects/events themselves. Specifically, MOVEMENT/TRANSPORTATION in *M4* is constrained by the lexical concept(s) associated with the word *water*. Even though *water* is taken as a natural kind term and the movement/transportation can be driven by some kind of natural force (e.g. current flow, the wind), a more intricate relationship is involved. The accessible paths available to understand the mental representation make the relationship open to more than one explanation. Water, in its natural kind form, can either be conceptualized as a resource/force to be harnessed or as a physical working environment to accommodate. Either situation sees human endeavors involved and highlights the physical attributes of water which are relevant and significant to achieve certain purposes. This seems to meet, at least partially, the criteria for one to identify a nominal concept, since water in ME sub corpus is conceptualized quite conventionally in the discipline by participants' purposefully picking out "features meeting particular descriptions" (Komatsu 1992: 513).

Conceptual dilemmas also arise with the lexical concepts associated with *electricity* and *temperature*. Electricity as a set of physical phenomena is associated with the presence and flow of electric charge which is the intrinsic property of some subatomic particles. Human can interfere into the producing process of electricity using energy sources (e.g. chemical combustion of fossil fuels, the kinetic energy of flowing water and the wind, nuclear fission, etc.) that create movement in a turbine or piston which rotates the rotor in a generator. The generator then utilizes electromagnetic induction to convert mechanical rotation into electrical current using a commutator. Electricity generation reveals a process of how people apply

knowledge acquired from observation and experiment on naturally occurring things to boost industrial production and economic development for human well-being. Similarly, temperature as an objective comparative (*A6.3 Comparing: variety*) measure (*N Numbers & measurement*) of hot or cold has a direct association with human sensory experience (*X3.3 Sensory: touch*) and is frequently referred to as a variable for temporary meaning construction in certain contextualized evaluative system. In other words, the word *temperature*, in spite of representing an objective reflection of a certain aspect in the physical properties of substance/object, is more often attributively used to help understand what it means by a temperature value to human needs/interests in various contexts.

There is thus serious limitation in Putman's conceptual categorization. Keil (1989) argued that rather than thinking in terms of a strict dichotomy between natural kind terms and nominal terms, one should recognize a continuum along which words that tend to be used referentially or attributively can be presented by degree. From such a perspective, a continuum of the lexical concepts involved in the 6 key semantic domains in ME sub corpus is proposed (Figure 11).

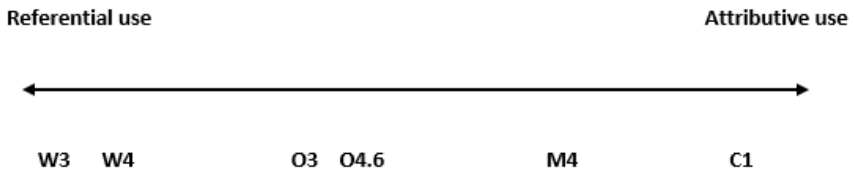


Figure 11 Conceptual continuum for the key semantic domains (ME)

The sub key domain *M4 Movement/transportation: Water*, by its very name, covers the PHYSICAL ENVIRONMENT and MEANS of the movement or transportation. In the USAS

tagging system, the domain is placed in parallel with *M3 Movement/transportation: Land* and *M5 Movement/transportation: Air* within the main semantic domain *M Movement, location, travel & transport*. The semantic categorization reveals people's application, in the process of their using language, of different sorts of MEANS to achieve the purpose of movement and transportation. *Water*, while may be classified as a *natural kind term* by Putnam (1975a, 1975b), is used here quite attributively (Donnellan 1977) since only some of its physical properties are taken as relevant and important (e.g. the property of being able to *flow* and having *buoyancy*) in the context. What is noteworthy is the fact that the participants do not simply select relevant properties of water for factual descriptions, but orient to taking them as variables and influential factors in dynamic interactions between human needs and the natural world. The experience-based schematizations of such interactions characterize the linguistic patterns as a representation of the cognitively structured mechanisms (Rumelhart 1980; Petruck 1996), as shown in the valence description of verb concepts in the [MOVEMENT/TRANSPORTATION BY WATER] frame (Table 19).

	Water	**Vessel**	**Human**
Flow	Subject {cause: natural kinetic force}	{effect: spontaneously moving/transported}	{role: apply/control}
Buoy	Subject {cause: natural upward force}	D-object {effect: floating}	(D-object {effect: floating}/ {role: apply/control})
Propel	{effect: accelerated behind rotating blades}	D-object {effect: moving forward}	Subject {cause: transmit rotational motion into thrust to push water backward}
Sail	on/in/through	(Subject/D-object) {effect: moving forward}	Subject
Anchor	*to the bed of*	(Subject/D-object) {effect: stopped}	Subject

Table 19 Valence description of verb concepts in the [MOVEMENT/TRANSPORTATION BY WATER] frame (M4)

Table 19 shows that being able to *flow* and *buoy* as the most basic physical properties of water in its natural form has direct effects on vessels, with human intervention to apply or control the natural forces. The interactional patterns between human and water determine what facets of the frame are to be activated for situated meaning-making. To *propel* a vessel, while involves using the physical properties of water (fluidity), emphasizes human interference to TRANSMIT rotational motion (of blades) into thrust to move the vessel forward by accelerating the water pushed backward. *Sail* and *anchor*, on the other hand, concern direct human involvement in CHANGING/MAINTAINING THE MOTION STATE of a vessel for different purposes. *Sail* describes changing a vessel's state from not moving to moving or the maintaining of its state of moving. *Anchor* describes changing a vessel's state from moving to not moving, or the maintaining of its state of not moving. In such sense, *sail* is used from a "not-moving" point of view while *anchor* is used from a "moving" point of view. The two words show different ways to schematize the world (Fillmore 1982a) not only in terms of the direction of changing motion state but also according to how people evaluate the attributes of water to marine activities and manipulate the interactional patterns accordingly. Specifically, people apply the kinetic force of water flow when they are sailing vessels while overcoming the force when they are anchoring vessels. In the former case the physical property of being able to *flow* meets human goals thus is evaluated as positive, while in the latter case such a property goes against human will thus is taken as negative.

To summarize, this section provides evidence of typical semantic frames evoked by particular linguistic forms across the key semantic domains in the three sub corpora. The observations have consequences for researchers who are interested in approaching knowledge representation in spoken academic discourse. The findings suggest that linguistic forms are not used in discourse by virtue of their context-independent meanings; rather, they are bound

by what the lexical concepts share in meaning representation. The shared meaning represented by their associative lexical concepts explains why and how the linguistic forms can be systematically grouped for meaning construction. The lexical relations between words (paradigmatic vs. syntagmatic) categorized in the field of *Lexical Semantics* (see Cruse 1986) are therefore insufficient to account for the complexity of speakers' lexical choices to make discourse coherent and accessible. By contrast, semantic domains as coherent structures of conceptual knowledge (Langacker 1987) have advantages over the paradigmatic-syntagmatic dichotomy to understand the mechanism of fine-grained portioning of meaning representation in discourse. The examples provided in this section show 1) that semantic domains "provide a way of carving out the scope of (lexical) concepts" (Cienki 2012: 14) relevant to identifying and characterizing patterns of their usage (Evans 2006); and 2) that linguistic forms grouped this way help to understand generic features of the target discourse.

5.3 Invoking interpretive frames: lexical concepts for knowledge activation

Lexical concepts constitute linguistically encoded conceptual knowledge that can be externalized via language (Evans 2006: 501). The coherent body of such conceptual knowledge is called a *frame* as a system of united linguistic choices related to schematic scenarios in motivating context (Fillmore 1982a). Invoking interpretive frames depends on how knowledge structures can be accessed through the path provided by lexical concepts to achieve certain communicative goals. This section focused on how the participants invoke specific interpretive frames for situated meaning-making to construct knowledge. Properties of lexical concepts associated with specific linguistic forms were examined to show how coherence is assigned to the talk and how knowledge structures are accessed through individual framings at the discursive level.

5.3.1 *Lexical profile and semantic tendency of lexical concepts*

A lexical concept by definition has a unique lexical profile to refer to general patterns of selection in terms of its grammatical, semantic, and collocational tendencies. (Evans 2006: 504). This is the unique "biometric identifier" of a lexical concept to define the unique "selectional requirements" of its usage (Evans 2006: 503). In his discussion of distinct lexical concepts encoded by two polysemous forms *time* and *fly*, Evans (2006) used single-sentence examples to illustrate the existence and characteristics of their lexical profiles (see also Evans 2004a, 2004b, 2005; Evans and Green 2006). This methodology, as suggested by Evans himself, is only to serve as a guide for further research rather than providing an established way of inquiry. He argues that it is an empirical question to examine a given lexical concept and suggests the use of corpus-based approaches to that end (Evans 2006). Corpus Linguistics methodology has advantages over observations of invented, isolated sentences, for the former works on real language use and examines patterns of lexis which might be overlooked without the assistance of computers. In this section, I presented two examples of distinct lexical concepts encoded by two words: *student-teacher* and *transcript*, by focusing on their lexical profiles in terms of semantic tendencies.

Student-teacher is a word calculated in *P1* in ALE sub corpus. This compound word is quite interesting even with a simple glance at its dictionary definition:

Student-teacher: a student who is learning/engaged in how to teach and practicing teaching for the first time

(Merriam-Webster Learner Dictionary)

A cognitive conflict is involved concerning how exactly the word is to be interpreted semantically. When *student* is viewed morphologically as the determiner for *teacher*, it seems to be used to prescribe a unique status of "being a teacher"; while according to the definition above, the core concept of the compound word is *student*, with *teacher* used to project a goal or a potential therefore prescribing a special kind of "being a student". So, is a *student-teacher* more precisely a *student* or a *teacher*? The confusion may not be related to the word order, since *teacher-student** seems to be an anomaly. The elusiveness of *teacher-student*[20] as a well-established compound word may reveal the conflict further beyond the syntactic level to the semantic level. If the second-part word in the compound has a retrospective tendency to denote some "condition" for the first-part word, such as the case for *student-teacher*, it seems that people do not usually say *teacher-student* because they do not usually think that *student* can determine conditions of being a *teacher*. In other words, people view the transformative pass between *teacher* and *student* in a one-way direction, relying on their schematic cognition shaped by the asymmetrical power distribution in the teacher-student relationship. In such sense, an appropriate explanation can be that *teacher* and *student* respectively provide descriptions for the same referent, the relationship of which is coined as appositional (see Burton-Roberts 1975; Blakemore 1996; Loock and O'Connor 2013) for semantic classification of compound words.

[20] A quick searching in BNC shows that *teacher-student* is sometimes used as a whole as a determiner for another noun, e.g. *teacher-student interaction*, *teacher-student ratio*, etc.

N	Concordance
1	mine because some of the teachers introduced me as a student teacher. So automatically the student says "oh well
2	. So automatically the student says "oh well it's just a student teacher" but then I was like "well ex= excuse me I
3	do you. Did you feel uncomfortable being identified as a student teacher+ Yeah. +or do you think that's kind of part
4	." I wouldn't have lied to them if they'd asked us are you a student teacher. I'd have been "well actually yeah I am."
5	way you're introduced as a colleague rather than oh just a student teacher. See we were introduced as colleagues
6	think maybe the odd some of them probably knew I was a student teacher. Em. Which didn't really matter. But I think
7	and that was nice to like be like so it's not just cos I'm a student teacher they're like that for everyone. But yeah. I
8	of a long placement never given out an "outstanding" for a student teacher as an overall grade cos he's like even
9	admitted it to them but told them straight off "yeah we're a student teacher" and came from that point of view that you
10	whereas with Dawn they just sort of stamped on you as a student teacher. Straight in like that was your identity in
11	got a bigger part to play in the school. Not just as a as a student teacher. You actually get integrated in the
12	. I don't think it was deliberate that he introduced me as a student teacher it just kind of came out Em. And it I mean
13	aren't you weren't concerned about being introduced as a student teacher yourself then perhaps don't realise that
14	. Yeah. Yeah. There were times that I forgot that I was a student teacher and just felt so included in it all and I just
15	as you start to feel more like a teacher as opposed to a student teacher who really doesn't know quite what they're
16	had. Mhm. Mm. So that that notion that although you are a student teacher that actually you you you're included in the
17	what have what does five weeks in a school offer me as a student teacher the chance to develop?" And what are the
18	in to the department. Because you've got um just another student teacher in with you and you catch things of each
19	colleagues but we eh like actually introduced ourselves as student teachers so we said we wo= o= were like really
20	you. Some schools have badges which identify people as student teachers or as associate teachers. And they say
21	Ofsted um criteria for satisfactory good and outstanding at student teacher level. I mean that was that was introduced
22	it is. Y'know | You know what the expectations are on student teachers. Um and I we I think I forget actually
23	them from the start for every class was the fact that we're student teachers we're from Newcastle University and
24	gonna be incidents but rather than just saying "oh you silly student teacher and y'know | you know I'll sort this out for
25	to say about if you like the way that the school offered the student teachers as a group um moments in time when
26	you were like part of something rather than just sort of the student teacher. Yeah. Yeah we had. Mhm. Mm. So that
27	you know you are treated as part of the team not just the student teacher you are like given given roles whereby

Figure 12 Concordance lines of *student-teacher(s)* (ALE)

The cognitive complexity of *student-teacher* at both the syntactic and semantic levels has a direct effect on how it is perceived as a discursive identity. Figure 12 shows the concordance lines of the lemma *student-teacher* (sorted at the L1 position). In most cases, student-teacher is viewed as a collective identity rather than used to refer to a certain person in the sub corpus, which is evident from its strong collocational association with the indefinite article *a* at the L1 positions. However, is the identity of a student-teacher perceived the same way by the participants as it is defined in the dictionary? To answer this question, I investigated all the 22 concordance lines of *student-teacher* and identified some lexical concepts associated with it mentioned by the participants (Table 20).

Concept category	Examples	Relation
Other actors	*mentor, other colleagues, partner(s)*	Authority to provide help and guidance
External description and evaluation	*really doesn't know quite what they're doing yet, hugely critical of what you have and haven't, associate, " oh you silly student teacher"*	Devaluation to a large extent
Internal struggles and expectation	*offer me… a chance to develop, realize that it might be of a concern to you* (to be introduced as a student teacher), *let them to buy in to your identity as a teacher, accept some of that authority, the label…the introduction can make a difference, change the level of confidence and identity, he'd never given out…an outstanding for a student teacher as an overall grade, it's not just cos I'm a student teacher, there were times I forgot I was a student teacher and just fell so included in it all*, he introduced me as a student teacher it just kind of came out, well it's just a student teacher, introduced as a colleague rather than just a student teacher, they just sort of stamped on you as a student teacher, treated any differently to another staff, like part of something/the team rather than just sort of the student teacher,*	A strong awareness of the "difference" in being a student-teacher, feeling marginalized and devalued in the institutions
External expectation	*gain more confidence, felt more like a teacher, be able to prove yourself, reflecting on your own experience, criteria for satisfactory good and outstanding, although you are a student teacher…you have a responsibility*, did you feel uncomfortable being identified as a student teacher?**	as opposed to a student-teacher

Table 20 Relevant concepts associated with *student-teacher* (ALE)

The word *student-teacher* evokes the [IDENTITY] frame and the distinct lexical concepts it encodes from the student-teachers' side and other institutional actors' side reveal a discursive divergence for interpretation. Nevertheless, what is shared by both sides is a strong negative representation of the identity when it is observed at the workplace.

The other word *transcript* is used in the MARG talk. In ALE sub corpus, *transcript* is semantically tagged into the domain *Q1.2 Paper documents and writing*. Conceptualized as a relational noun, *transcript* is expected to have semantic associations with some other words in

a specific way. However, an investigation of its collocates does not reveal much about any coherent relational patterns.

N	Word	With	Relation	Set	Texts	Total	Total Left	Total Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	TRANSCRIPTION	transcription	30.194		2	6	0	0						6					
2	TRANSCRIPT	transcript	26.372		3	31	0	0						31					
3	TRANSCRIPTS	transcripts	25.456		2	12	1	1			1			10			1		
4	WHICH	transcript	17.187		3	3	1	2	1						2				
5	BACK	transcript	12.895		2	2	1	1			1				1				
6	ON	transcription	12.669		2	3	3	0		1		1	1						
7	AT	transcript	12.121		3	8	6	2	1		1	4			1	1			
8	STARTED	transcription	12.014		1	2	1	1		1									1
9	JUST	transcript	11.534		3	5	5	0	2	1	2								
10	FIRST	transcript	10.853		2	2	0	2							1	1			
11	NOT	transcript	10.813		3	4	2	2	1	1						1	1		
12	LOOK	transcript	10.604		2	2	2	0			2								
13	WHEN	transcript	9.799		2	2	1	1				1							1
14	BETWEEN	transcripts	8.547		1	2	2	0	1			1							
15	MY	transcription	7.762		1	2	2	0			1	1							
16	WORD	transcripts	7.675		1	2	2	0	1		1								
17	SAY	transcript	7.644		2	2	2	0			2								
18	WITH	transcript	7.333		2	2	2	0				1	1						
19	CAN	transcript	7.100		2	3	2	1	1	1									1
20	PUT	transcripts	7.046		1	2	2	0	1			1							

Table 21 Top 20 collocates of *transcript* in MARG talk (ALE)

Table 21 shows the top 20 collocates of *transcript* in MARG talk. Most of the collocates cluster at the centre of the span. Those which are not at the centre positions do not seem to reveal coherent semantic patterns since the majority of them are functional rather than content words. This indicates that the semantic associations (if there were any) between *transcript* and other words might not be readily accessible from a statistical perspective. The low frequencies of the collocates, in addition, indicate that little particularity or generality can be claimed. This means that it is quite difficult (if not impossible) to discover what semantic features are shared by the collocates of *transcript*, namely, its *semantic preference* (Louw 1993; Hunston 1995; Bednarek 2008). Since knowledge of semantic preference is probably shared among speakers of a certain speech community (Partington 2004: 152), the occurrence of a strong semantic preference of lexical items has a great potential to "prime" (Partington 2004; Hoey 2005) specific expectations in interaction through cumulative effects of encounters (Bednarek 2008). However, in spite of the elusiveness of salient patterns of its semantic preference, the

word *transcript* is indeed a term which is heavily loaded with shared knowledge among CA analysts thus sees no difficulty at all in priming discursive expectations. To further explore if there might be any pattern of relation emerging from context, I examined the concordance lines of the lemma *transcript* (Figure 13).

aren't they? Actually I was planning to bring another	transcript	here today but erm there was a there was
if you look at erm Goodwin's transcripts. If you look at	transcripts	they have a a line. And so it of. Mm. yeah
within the transcripts? Or are they going to be	transcripts?	It will be difficult to put pi= pictures. So
. I've seen people er separate them between different	transcripts	and same transcripts by two columns. Like
mentioned as limitations when they are doing	transcripts	that this student cannot be seen because
first could you find any problems on my descript= er	transcription?	Er the spelling of the word bonkers. Er.
it but if you if I'm going to put a comprehensive erm	transcript	I thought maybe I can do it this way. But I
I first ma= to moved to Word from Word Perfect for	transcripts	putting+ Mm. +the numbers down the side.
techniques then you have to look at everything there is	transcripted.	. I know. It's a lot. And that's going to
that? Well yeah cos I thought like there's then loads of	transcript.	It will be there when. Oh cos it yeah.
as a list when. I know. Mm. Er just while we're on	transcription	as well. More like erm erm I don't know
it little bit better if you're not looking at an organised	transcript.	But let's just see how it goes. It lasts two
myself. Any questions before we start? Um regarding	transcript	conventions. I think that's more or less
well. I mean that's like if you look at erm Goodwin's	transcripts.	If you look at transcripts they have a a
separate them between different transcripts and same	transcripts	by two columns. Like one for non verbal
Paul yeah. Yeah. P A U. Have you got any spare	transcripts?	Mm? Have you got any spare scripts? .
I said before I came here I learned analysed my st=	transcription	data and I checked stepwise topic
lines as you mentioned like erm but then in in that	transcript	I couldn't capture when the mutual gaze
you said at the beginning. You know don't read the	transcript	first listen to it. Mm. Because if you know if
to say as you can see background coughing the	transcript	is not the finished article. recording is
. Oh. Just one more thing that I would say about the	transcript	as well is that erm. At line twenty-five the
he didn't know that this is a call line just looking at the	transcript	as as a talk as it is we could think that
let me just go back. Alright I'm going to give you the	transcript	as uh. background rustling If you take um
the call to begin with. I don't want you to look at the	transcript	because the reason for that I want to do
context probably and make some comments about the	transcript	and then. We'll just go just let's just let's go
need some time to think? Okay. Well is asking the	transcript	erm. Yeah. Is it a formatting thing? Or is it
through twenty-one when I look at it just at the	transcript	I'm not quite sure how that works in terms
cannot be captured and cannot be reflected in the	transcript.	Erm the girl Kim is here so we cannot see
on that page. Yeah. I thought just again just on the	transcript	. I thought thir= seventy-eight. Seventy-nine.
the transcription as we as we started talking about the	transcription.	Uh-huh. Well this is something now I'm
gonna get it checked. Erm as for the beginning of the	transcription	as we as we started talking about the
screenshots. So is it going to be embodied within the	transcripts?	Or are they going to be transcripts? It will
in between the analysis. Not necessarily within the	transcript.	If I can with arrows. Mm-mm. Yeah. I can I
and see it again. Thanks. You can have a look at the	transcript	system at the ending . Uh-huh. It has
mean just looking across here. If you put if you put the	transcripts	like this. There's a lot more interactional

Figure 13 Concordance lines of *transcript* in MARG talk (ALE)

Again, few salient semantic patterns can be identified from the concordance lines. The definite article *the* dominates the L1 positions of the lemma, indicating that a transcript is most frequently referred to as a physical object by the participants in their talk. This necessitates a shift of focus to the pragmatic side of the word upon which the *semantic prosody* (Louw 1993; Xiao and McEnery 2006; Bednarek 2008) can be investigated since its "normal semantic values are not necessarily relevant" (Sinclair 1996: 34) in the context.

According to Sinclair (1996), semantic prosody is "attitudinal" (p. 34) to express the function of the target linguistic item collaboratively achieves with its textual surroundings.

Central to semantic prosody is human evaluation. The pragmatics fundamental to a linguistic choice in context reflect the evaluative meaning (Stubbs 1995; Bublitz 1995; Hunston 1995; Hunston 2002; Partington 1998) derived from the judgment of the value of discourse. Discussions and debates in Corpus Linguistics generate claims concerning how such evaluation can be categorized: the semantic prosody of a linguistic choice can be positive, negative or neutral. However, as argued by Bednarek (2008), the re-occurrence of a linguistic item within a positive/negative context does not automatically indicate the presence of a corresponding connotation in its meaning. In addition, the so-called neutral prosody is far from being unproblematic since it is often difficult to confirm the subtle balance grounded in limited datasets (see Stubbs 1995; Channell 2000 for different views). I argue that the semantic prosody of *transcript* cannot be simply identified as "positive" or "negative". A hierarchical structure exists within which the participants co-select lexical concepts to invoke interpretive frames in terms of how evaluative meanings of *transcript* are conveyed and negotiated to construct knowledge.

The bottom of the hierarchy addresses the participants' shared knowledge of the limitation of transcripts-in-use. The [ASSESSING] frame is invoked to evaluate transcripts as an end-product and a facilitative tool. The collective awareness is believed to rise from their familiarity with CA theories and a cognitive consensus formed through the long-term training of relevant techniques. The expanded talk around the lemma *transcript* shows a strong evaluative sense with negative connotations. For example, the participants frequently mentioned that transcripts "*couldn't capture and reflect*" certain aspects of communication,

that some features are missing in transcripts since speakers "*cannot be seen*" due to technical deficiency (e.g. camera position), that "*a comprehensive transcript*" is needed, that the transcriber deliberately initiates problem-detecting moves among the participants on his/her transcripts, and that there is something that "*we cannot say according to transcript*", etc. The lexical patterns identified at this level correspond to Partington's (2004) observation that the overall meaning of some lexical patterns can be highly favourable despite that the words co-selected with the target item have unfavourable connotations. The negative lexical concepts associated with *transcript* show the participants' critical attitude which is believed to be a good practice in CA research.

In the middle of the hierarchy lies the participants' endeavour to optimize the accuracy of their transcripts. Relevant lexical concepts are mentioned to invoke the [CAUSE TO MAKE PROGRESS] frame. Issues discussed around the concept of transcript focus on a series of challenges at the action level, including how to identify topic shifts accurately, how to represent multimodal features (e.g. mutual gaze, head movement, gestures) appropriately, how to balance the amount of information in a transcript, how to ensure spelling consistency, how to standardize the formatting for readers' convenience, etc. Contextualized, directive orientations and personalized framings (Csomay 2005: 260-261) are found to be discursively salient in the talk. Language use at this level also reveals the participants' ambivalent attitude to the existing transcribing conventions and the flexibility of their application in their own transcription. The ambivalence shows a less visible resistance to the authoritative, well-established conventions concerning how communicative realities should be represented in the research practice.

At the top of the hierarchy is the participants' intensive focus on the interactional structures represented in transcripts and how the analysis of such structures potentially leads to interesting findings. The [SCRUTINY] frame is invoked to show the analytical nature of talk-in-interaction. What is salient at this level is the participants' repeatedly explicit referring to CA terminologies that define and describe structural characteristics of interaction. However, a discursive orientation can be identified at this level that the participants go beyond the immediate context of what is textually limited by the transcript. There is a conceptual transition from the structure to the content and then to the implication of the data. The transition is procedurally facilitated and collaboratively achieved.

To summarize, the observation of the lexical concepts associated with *student-teacher* and *transcript* reveals that the lexical concepts are co-selected and organized in patterned ways to invoked interpretive frames for certain communicative purposes. For *student-teacher* as an identity, both the student-teachers and their senior counterparts associate with it negatively evaluative lexical concepts. The selectional patterns, however, are constructed in different ways to activate different frames of conceptual knowledge. Student-teachers draw on the mental representation of "institutional equality", while their senior counterparts emphasize "professional qualification". This results in conflicting framings of the identity of student-teachers. The lexical concepts encoded by *transcript*, on the other hand, reveal how institutional constraints have an impact on the construction of conceptual knowledge to achieve consensus and coherence in a certain discourse community. The network of the interpretive frames invoked by the participants also reveals that frame shift and integration can relate to knowledge of particular action procedures.

5.3.2 *Semantic value of lexical concepts*

This section focuses on the semantic value of lexical concepts and how certain dimensions of semantic value are responsible for frame invoking to interpret the target discourse. According to Evans (2006), the semantic value associated with a lexical concept has at least five dimensions:

- Semantic potential;
- Encapsulation;
- Relational vs. non-relational;
- Temporal structure; and
- Referentiality

While the first relates to non-linguistic knowledge, the other four are unique to the linguistic knowledge associated with a given lexical concept (Evans 2006). This section provides a series of observations concerning how different aspects of the semantic value of lexical concepts activate situated interpretation of the participants' small group talk. Note that not all of the five dimensions cited above will be addressed or equally addressed according to the features of NUCASE data. The observations will show that the participants' language use for meaning construction tends to partially activate a frame of conceptual knowledge thus highlights certain facet(s) of the knowledge structure (Evans 2006).

The interpretation of the participants' language use, on the other hand, is affected by the inquiry method used in this section which is characteristically distinct from what has been widely accepted in semantics research. Rather than examining a group of isolated sentences containing a particular linguistic form, I investigated instances of the form retrieved from the

target discourse to let any semantic regularity emerge from context. The corpus-based approach to meaning has an interpretive consequence on the selection of lexical concepts to construct knowledge. For example, Evans (2006) discussed how linguistic and extra-linguistic contexts would give rise to the selection of a particular lexical concept associated with a polysemous form (e.g. *hike*, *bar*). For a corpus inquiry, however, such openness to selection is to a large extent made obsolete because the context has already been there to limit the scope of selection. This corresponds to Croft's (1993) argument that one needs to know what the entire utterance means in order to know the meanings its respective words take on.

For the linguistics-related knowledge represented by lexical concepts, I shall start by addressing the "relational vs. non-relational" dimension of semantic value. For example, the adjective *academic* in ALE sub corpus is found to be associated with lexical concepts in five dimensions:

- Level of learning associated with an institution (e.g. "*academic* year/talk/freedom");
- Level of formality related to performance or work (e.g. "*academic* excellence");
- Level of abstraction (e.g. "*academic* questions/subjects/concern");
- Level of significance or usefulness (e.g. "a question which is *purely academic*");
- Level of conventionality (e.g. "*academic* painting")

Merriam-Webster's Collegiate Dictionary 2.2 (2005)

As one shall see, all the dimensions above are relatively open-ended, along which no point can be identified as the most typical to represent a prototype (Gärdenfors 2014). In addition, the dimensions can be associated with one another to different degrees. For example, the first three dimensions: "level of learning", "level of formality" and "level of abstraction" are more

closely associated with each other; while the last two dimensions: "level of significance or usefulness" and "level of conventionality" show weaker associations with the first three dimensions and with each other. The first three dimensions are found to be associated with an *institution* which can be taken as *academic* in terms of its goals, activities, and representation. With the difficulty locating a prototype to represent what is the "most *academic*", the relational nature of the word is to be better examined through comparisons between instances at a particular level of meaning.

- Level of learning: goals and evaluation of the goals of the institution;
- Level of formality: activities and evaluation of the activities in the institution;
- Level of abstraction: representation and evaluation of the representation in the institution.

The observation reveals that only some of the frame facets are activated for situated meaning construction. As shown above, the form *academic* in NUCASE activates teaching and learning experience in an institution (e.g. not a writing style).

Lexical concepts associated with nouns are taken by Evans (2006) as non-relational. Non-relational nouns are classified into entity categories by Gentner and Kurtz (2005). Lexical concepts associated with them are perceptually given to refer to objects and meaning-making is accordingly correspondence-driven. Here is an example. In the MARG talk in ALE sub corpus, the entrance into a talk scenario is marked by the participants' explicitly referring to the page/line number in the transcript at hand. *Line* and *page* can be perceived as non-relational nouns (e.g. *O4.4 Shape; Q1.2 Paper documents and writing*) since the fundamental pragmatics to them are references to objects based on the knowledge acquired by individuals

independent of concrete discursive practice. However, *page* is used slightly different from *line* in the talk. While the latter evokes a salient pattern of "*line* + cardinal number/the (+ ordinal number) + *line*" to refer to one or more certain lines in the transcript; the former, in addition to the textually referential use (Figure 15), is also used to talk about conceptual quantification of the object (*page* as a unit of document) in CA transcription and analysing activities (Figure 16).

what she's talking about? I'm trying. On this last page. I don't know. "I can hear the frustration in
er three one seven. Three one seven. Yeah. Last page. Er. Three one seven? Ah no no no.
. Yeah. Or would you prefer to look at the last page? Okay. Which one? First page. First page.
because if you look at what the guy says on the next page the technician says on pa= on line forty-nine if
or+ Yeah. +half a page or something. Yes one. Page one. Erm. No. Line thirty-one to line thirty-four
part you'd like us to focus on? Yeah. Maybe one page or+ Yeah. +half a page or something. Yes
languages and cultures? Again on that same page I think. I'm not one hundred percent sure of
you. So then if you look at the to look at the second page and just doing a a really non CA analysis by
sorry. Erm on line one five four while we're on that page. I think there is a bad connection and it it
yeah. Thirty-two. Thirty-five. Forty-four. Just on that page. Yeah. I thought just again just on the
. Mm. So we could say that for the first third of the page the student is talking about the benefits of
see you can just it actually when you just cross the page. If you if you look at what's going on in the
conflating two things when we look at this page because there is there is a sequence when

look at. Shall we pick one either I think the first page or the last? Yeah. I'd say the first pages.
pages. Okay. Shall we just look at the whole first page and then. Okay. For ten minutes. Hmm. Yeah
don't we just break it down maybe look at the first page? Okay. First. Five second pause Shall I say a
to look at the last page? Okay. Which one? First page. First page. Nine minutes thirty seconds
the last page? Okay. Which one? First page. First page. Nine minutes thirty seconds pause . your
the bit. Five minute silence Can we watch that first page again please Andre ? Yep. Silence, then

Figure 15 *page*_textual reference

focus on? Yeah. Maybe one page or+ Yeah. +half a page or something. Yes one. Page one. Erm. No. +he might want something to+ +draw a line across a page or there's a whole number of different things.

he's making it more dramatic he's gonna put a four page ad in. ratchetting+ Yeah. Yeah. Yeah. +up the
that. Line fifty-four he's said he's gonna put a four page ad in the ha ha laughing in the paper for six
said he's gonna put a four page ad in the + Oh four page right. +in the+ +paper for six grand. Yeah
. +it becomes oh he's said he's gonna put a four page ad in the + Oh four page right. +in the+
a half page that's interesting. A half page. Not a full page. laughter That's kind of interesting. Why the
to spend six and a half grand and take out a half page ad in the financial and explain the shit that
kind of interesting. Why the hell would it be a half page? However but+ I'm I just noticed that. + at
one of my points here is that it's kind of the half page and the six and a half grand and the financial
. Yeah. Cos he knows exactly how much+ . +a half page ad+ Yeah. + off the top of his head. Yeah+
between the and that. Oh. Explain. Stop . Take half page ad in financial review and explain so sorry
company and s= I can spend six half grand in a half page and then I can't even get one single piece
to spend six and a half grand and take out a half page ad in the financial and explain the shit that
to spend six and a half grand and take out a half page ad in the financial and explain the shit that
in in terms of his threat which is this what is half page uh a half page that's interesting. A half page.
of his threat which is this what is half page uh a half page that's interesting. A half page. Not a full page
six grand. Yeah yeah that's interesting. Yeah a half page. So he's making it more dramatic he's gonna
is half page uh a half page that's interesting. A half page. Not a full page. laughter That's kind of
exactly what he's doing. Yeah. That's why it's a half page ad. Yeah. Cos he knows exactly how much+ .
to spend six and a half grand and take out a half page ad in the financial and explain the shit that

Figure 14 *page*_conceptual quantification (ALE)

However, Evans' (2006) over-simplified categorization of relation representation according to a word's part-of-speech has long been challenged. Some nouns are found to be relational indeed, for which the standard treatments of nominal semantics do not apply since what it encodes are more than individual lexical concepts but a pattern to describe how they are related to each other. Relational nouns are "semantically unsaturated" (de Bruin and Scha 1998: 25; Asmuth and Gentner 2005: 163) and there are significant differences between syntactic structures connecting relational nouns to their functional/values and non-relational nouns.

Relational nouns have some commonalities with verbs, in that their meanings are "centred around extrinsic relations with other concepts" (Asmuth and Gentner 2005: 163). For example, the noun *transcript* is more appropriately to be perceived as a relational noun. The relational denotation can be partially traced from the verb form *transcribe,* an action constituting the

relationship between the verbal/non-verbal reality to be transcribed, the transcriber and the transcript as a representation of the reality (Figure 17).

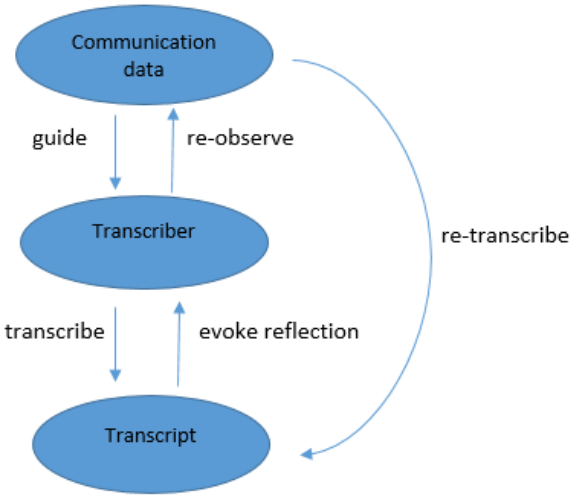


Figure 15 Cyclic process for CA transcribing

The understanding of relational nouns usually occur later in language acquisition process than non-relational nouns (Gentner 1982; Gentner and Boroditsky 2001; Genter and Kurtz 2005) and is gradually enhanced through experiential practice. For example, the noun *transcript* is typically understood first as an object reference (e.g. *I'm going to give you the transcript*) before evolving to a more relational interpretation (see Gentner and Rattermann 1991; Waxman and Hall 1993).

The status of being "relational" of nouns is not always so straightforward that it can be easily identified. This applies particularly to nouns which are semantically categorized into natural kind terms (Putnam 1975a, 1975b), such as *shore* and *coast* in ME sub corpus. Both *shore* and *coast* are semantically associated with the lexical concepts LAND and WATER: they

mean land along the edge of a body of water. Cross many languages, the two words are interchangeably translatable thus are taken as entity nouns. Fillmore (1982a) in his work described the details of the difference between these two words as a bit tricky, though they both evoke the [BOUNDARY] frame. The shore is the boundary between land and water from the water's point of view while the coast is the boundary between land and water from the land's point of view (Fillmore 1982a: 383). In such sense, a trip which takes six hours "from coast to coast" is a trip across a land, while a journey "from shore to shore" spatially covers a body of water. Similarly, traveling "to the shore" is about a sea journey while a trip "to the coast" means one on land. Strong collocational associations are identified in ME sub corpus between the word *shore* and expressions denoting concepts of LOCATION AND DIRECTION (*M6 Location and Direction*) (e.g. *on*, *nearer to*, *back to*, *off*, *into*, *from*) (Figure 18).

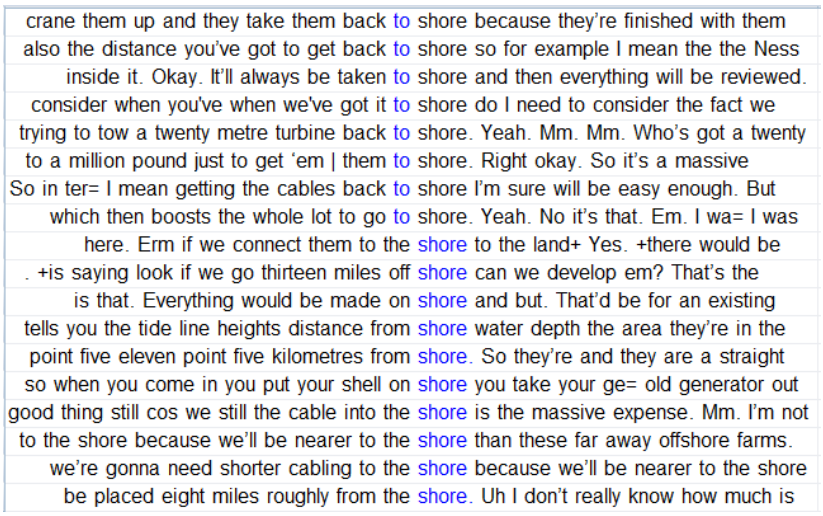

crane them up and they take them back to shore because they're finished with them
also the distance you've got to get back to shore so for example I mean the the Ness
inside it. Okay. It'll always be taken to shore and then everything will be reviewed.
consider when you've when we've got it to shore do I need to consider the fact we
trying to tow a twenty metre turbine back to shore. Yeah. Mm. Mm. Who's got a twenty
to a million pound just to get 'em | them to shore. Right okay. So it's a massive
So in ter= I mean getting the cables back to shore I'm sure will be easy enough. But
which then boosts the whole lot to go to shore. Yeah. No it's that. Em. I wa= I was
here. Erm if we connect them to the shore to the land+ Yes. +there would be
. +is saying look if we go thirteen miles off shore can we develop em? That's the
is that. Everything would be made on shore and but. That'd be for an existing
tells you the tide line heights distance from shore water depth the area they're in the
point five eleven point five kilometres from shore. So they're and they are a straight
so when you come in you put your shell on shore you take your ge= old generator out
good thing still cos we still the cable into the shore is the massive expense. Mm. I'm not
to the shore because we'll be nearer to the shore than these far away offshore farms.
we're gonna need shorter cabling to the shore because we'll be nearer to the shore
be placed eight miles roughly from the shore. Uh I don't really know how much is

Figure 16 Concordance lines of *shore* (ME)

The collocational patterns show that the participants talk about things (which are supposed to be) "in the water (sea)" when they use the word *shore*, which is consistent with what is called

by Fillmore (1982a) “from the water’s point of view” (p. 383). However, a closer look at the concordance lines reveals a textual preference in the participants’ talk that they invoke the [DISTANCE] frame to assign coherence to their talk of *wind farms* in the water. Distance is considered as relevant and significant because 1) it is a principle parameter for the selection and evaluation of the location for a facility (e.g. *eleven point five kilometers from shore, nearer to the shore, thirteen miles off shore*); 2) it is directly related to the convenience and cost of the maintenance of the facility since some of the work needs to be done on land (e.g. *getting the cables back to the shore, to tow a twenty meter turbine back to the shore, connect them to the shore*); and 3) it involves political considerations since the installation of facilities in the sea may be beyond the boundary of one country’s territorial waters.

The word *coast*, in its natural kind form, is attributively perceived as part of a country’s territory to be guarded, the security of which ensures the country’s sovereignty integrity. However, it is used in a much more limited manner in ME sub corpus. While collocating with more concrete lexical concepts such as *procedure*, *fleet*, *iceberg*, *equipment*, *vessels* (Table 22), *coast* is overwhelmingly used in the expression *the US Coast Guard*[21] (28 hits) in only one small group talk session. When used as part of a proper noun to refer to a unique entity, the feature of being “naturally existing/occurring” of *coast* has been diminished to a large extent while its attributive connotation is highlighted in the context.

[21] The United States Coast Guard (USCG) is a branch of the United States Armed Forces and one of the country's seven uniformed services.

N	Word	With	Relation	Set	Texts	Total	Total Left	Total Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	COAST	coast	62.17		5	35	0	0						35					
2	LINE	coast	18.92		2	2	1	1		1					1				
3	NORTHEAST	coast	15.28		1	2	2	0				1	1						
4	HOMELAND	coast	14.23		1	2	1	1	1										1
5	PART	coast	13.64		2	3	2	1		2							1		
6	PROCEDURES	coast	12.03		1	2	0	2									2		
7	FLEET	coast	11.55		1	4	1	3		1						1	2		
8	ICEBREAKING	coast	11.55		1	2	0	2								2			
9	REGULATIONS	coast	11.17		1	2	0	2								1			1
10	ON	coast	10.77		3	3	3	0			3								
11	EQUIPMENT	coast	9.66		1	2	1	1				1				1			
12	GUARD	coast	9.43		1	28	0	28							28				
13	WHICH	coast	9.25		2	3	2	1	2										1
14	VESSELS	coast	8.18		1	2	0	2								2			
15	NEEDS	coast	7.13		1	2	0	2								2			

Table 22 Top 15 collocates (L5 to R5) of *coast* (ME)

I shall now turn to another two dimensions of semantic value of lexical concepts: encapsulation and referentiality. The former means that lexical concepts serve to encapsulate "often complex and informationally diffuse ideas" (Evans 2006: 509) and the latter means lexical concepts refer to "entities of different sorts" (p. 510). The following example shows how tricky it could be to interpret the meaning of utterances when the complexity of an idea encapsulated by some lexical concepts is intensified by the referentiality of other lexical concepts co-selected in the discourse.

Teacher: *If you have a bad conscience, how do you feel?*

Student: *Bad (with laughter).*

The utterances above are retrieved from the language data the MARG participants collaboratively analysed in a small group talk session. The data is an audio-recorded classroom interaction scenario of teaching English as second language (L2) in a public school. The entrance topic in the teacher's instruction was "social conscience". The question above asked by the teacher was aimed to guide the students to think about the topic. The laughter uttered with the student's answer, however, indicates somehow that the teacher's question is a

strange one. With no additional contextual information provided, the question seems quite difficult to be answered (if not completely unanswerable). The difficulty arises concerning how to interpret the semantic value of the lexical concepts associated with *bad conscience*.

Conscience as a complex idea is encapsulated by lexical concepts associated with the [HUMAN SENSE] frame and the [MORALITY] frame. It is about the relation between people's mental judgment and their own conduct or intentions. In such sense, the use of the word *bad* to modify *conscience* is quite confusing since it is unclear the quality of which side involved in the relation *bad* actually refers to. The teacher's utterance is, as a result, open to multiple interpretations which would lead to dramatically different consequences:

1) *The incapability interpretation*. The speaker assumes that the hearer is incapable of telling properly what is right and wrong against a received social norm or as what other would do in similar situations. Negative moral evaluation is associated with the use of *bad*, showing a criticizing attitude of the speaker. In such sense, one may expect that the follow-up question (*How do you feel?*) might not receive any valid answer based on the assumption that the hearer may have problems "feeling" the "right" way.
2) *The unethicality interpretation*. The speaker assumes that the hearer is capable of telling whether his/her conduct or intentions is/can be wrong thus feels bad about it. The use of *bad* is thus for relatively neutral evaluation. In such sense, the follow-up question may be a valid one to induce more thought or reflection upon what social conscience means to individuals with outlooks that conform to accepted social norms.

The first interpretation is more closely associated with the [HUMAN SENSE] frame and the lexical concepts encoded by *bad* can be understood as similar to those encoded by *insufficient* or *disabled*. The second interpretation is more closely associated with the [MORALITY] frame and the lexical concepts encoded by *bad* can be understood as similar to those encoded by *guilty* or *troubled*. The observation has serious implications for the evaluation of the legitimacy and effectiveness of teacher instruction in L2 learning practice. Although it is sometimes inevitable for the teacher to introduce topics that are conceptually difficult and complex to talk about in class, he/she is responsible for providing sufficient contextual information to guide students to make sense of/with the topics. The observation has also highlighted the necessity of integrating specific pedagogical goals with students' established knowledge networks; namely, invoking proper frames to help students assign coherence to the discursive practice they are engaged in.

Finally, I shall address what has been perceived as the most important aspect of semantic value of lexical concepts which accounts for the protean word meaning in language use: a lexical concept possesses a semantic potential (Evans 2006: 509). Semantic potential gives rises to "an information characterization" with lexical concepts "providing access to conceptual knowledge structures" (Evans 2006: 509). The path of access determines what facet(s) of a knowledge structure is/are to be activated for situated meaning construction through communicative language use. Observation of the NUCASE data reveals evidence of such partial activation in the participants' frame applications to construct knowledge in local contexts.

Words' collocational behaviours shed light on how a semantic potential of lexical concepts provides a unique path of access to a conceptual knowledge structure. For example, the word

propeller in ME sub corpus has the 2nd highest frequency value (108 hits in singular form, 33 hits in plural form, 141 in total in 7 talk sessions) in *M4* and represents most explicitly how human interact with water by technically pushing it back to the rotating blades and accelerating it to generate forward thrust. The word *design* shows the strongest collocational association with *propeller* with functional words excluded from the top 10 collocate list (Table 23).

N	Word	With	Relation	Set	Texts	Total	Total Left	Total Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	PROPELLER	propeller	114.89		12	109	0	0						109					
2	THE	propeller	8.51		12	77	57	20	4	10	6	4	33			5	7	4	4
3	DESIGN	propeller	37.31		7	34	8	26	1	1		5	1		25	1			
4	A	propeller	10.40		8	34	27	7	3	1	9	7	7		1		3	2	1
5	AND	propeller	8.44		8	26	12	14	2	6	2	2			1	5	1	3	4
6	TO	propeller	8.38		8	20	10	10	1	2	2	5			6	2	1	1	
7	YOU	propeller	0.78		4	18	15	3	4	3	7	1			1		1		1
8	IS	propeller	9.22		6	16	5	11		2	1	1	1		6	4			1
9	WE	propeller	4.45		6	15	11	4	5	5	1					1	2		1
10	YOUR	propeller	4.74		2	13	8	5		1			7			1	1	2	1

Table 23 Top 10 collocates of *propeller* (ME)

The word *design* in the FrameNet lexical unit index is associated with the [COMING UP WITH] frame. The frame has to do with a person creating a new intellectual entity, an idea which is often conceptual in nature. Mental effort to figure out something out rather than the impact of the idea is emphasized in the interpretation of the frame. On the other hand, the word *design* is tagged by the USAS semantic tag set into Domain C1 *Arts and crafts* and Domain X7 *Wanting, planning, choosing*. This means that both the intellectual activities to design something and the impact/result of the design product are emphasized in terms of the word's semantic potential.

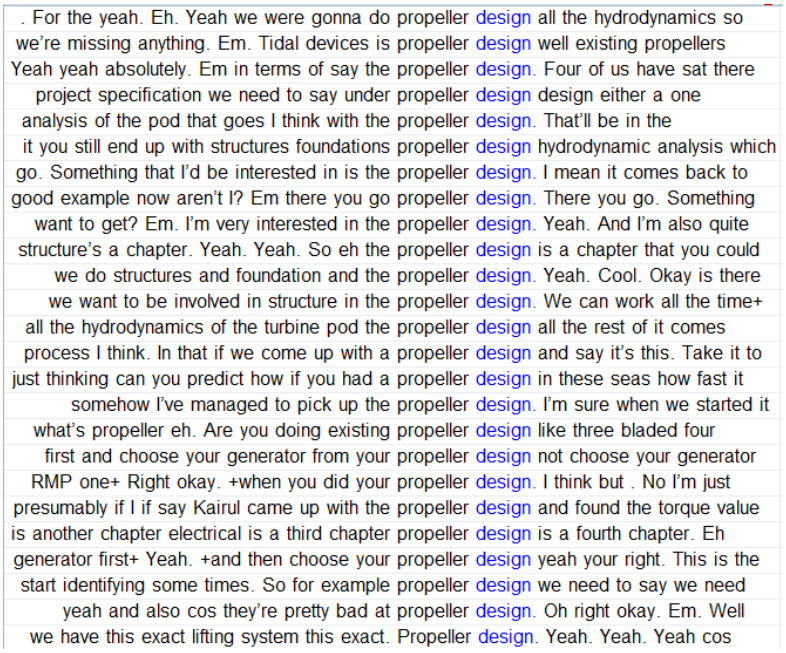


Figure 17 Concordance lines of *propeller design* (ME)

Linguistic forms encoding distinct lexical concepts are found to be co-selected with *propeller design* as shown in the concordance lines (Figure 19). The lexical concepts are either associated with intellectual/psychological activities to design a propeller (e.g. *very interested in/came up with/think with/choose your/gonna do/pick up the propeller design*, etc.) or the impact of the propeller design (e.g. *existing propeller design, propeller design is a chapter, structure in the propeller design, choose your generator from your propeller design, under propeller design*, etc.). The selectional patterns can be identified from the top collocates of propeller design (Table 24).

N	Word	With	Relation	Set	Texts	Total	Total Left	Total Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	UP	propeller design	18.30		4	4	4	0		1	2	1							
2	UNDER	propeller design	15.48		2	2	1	1					1						1
3	BLADED	propeller design	9.82		1	2	0	2									1		1
4	WITH	propeller design	7.77		3	4	4	0			1	3							
5	INTERESTED	propeller design	6.62		1	2	2	0		1	1								
6	CHOOSE	propeller design	6.31		1	3	2	1	1			1				1			
7	EXISTING	propeller design	5.73		1	2	1	1					1			1			
8	SAY	propeller design	5.02		2	4	2	2				2				1		1	
9	STRUCTURES	propeller design	4.58		1	2	2	0	1			1							
10	CHAPTER	propeller design	4.46		1	3	1	2					1				1	1	
11	GENERATOR	propeller design	4.28		1	2	1	1			1							1	
12	LIKE	propeller design	3.08		2	2	0	2							1				1

Table 24 Top 12 collocates of *propeller design* (ME)

Examination of collocational associations can also reveal subtle differences between forms derived from the same lemma concerning how a semantic potential provides access to different facets of a frame. For example, the participants in ME sub corpus tend to use the two words *electrical* and *electric* interchangeably sometimes. Both of them are associated with the [ELECTRICITY] frame. In most cases, however, the two words show distinct collocational tendencies. Basically, *electrical* is used to refer to things that are related to or concerned with electricity in a relatively general term. For example, *an electrical chapter/side/part* are associated with the lexical concept of ILLUSTRATION related to written work (e.g. in a thesis or report), while *an electrical modeling* is associated with the lexical concept of DEMONSTRATION related to visual aims (e.g. experiment presentation or graphs). *Electrical work* encodes lexical concepts such as GENERATION, SUPPLEMENT and DISTRIBUTION of electricity in the industry, while *electrical connection* encodes the lexical concept of CIRCUIT. By contrast, *electric* is used with entity categories, the collocational patterns of which indicate a direct association between electricity and objects. For example, *an electric (podded) propulsion/motor* is associated with the lexical concept of POWER in which electricity is the energy source to produce the propulsive force, while *electric current* is associated with the lexical concept of FLOW that represents the physical properties of electricity.

Ranking	Collocates of *electrical*	Collocates of *electric*
1	Engineering	an
2	equipment	field
3	appliances	Nuclear
4	Mechanical	Shock
5	activity	Co
6	goods	fire
7	engineers	light
8	conductivity	guitar
9	Engineer	heater
10	impulses	shocks
11	an	Matsushita
12	stimulation	cooker
13	Electronic	Oki
14	Amalgamated	drill
15	signals	current
16	wiring	Gas
17	signal	General
18	and	Electric
19	energy	kettle
20	fault	motor
21	electronics	Mitsubishi
22	circuit	guitars
23	Institution	power
24	appliance	windows
25	power	charge
26	properties	heaters
27	retailer	blanket
28	plumbing	motors
29	circuits	fires
30	Industries	sunroof

Table 25 Top 30 collocates of *electrical* and *electric* in BNC (CQP web)

The differences in the semantic preferences of *electrical* and *electric* are not unique to NUCASE data. Table 25 above shows the top 30 collocates of the two words in BNC, indicating that the observation derived from a small data sample can be consistent with what can be found in a much larger corpus. From the view of Frame Semantics, the textual preferences reveal how participants construe the features and functions of the same thing in different ways, which is reflected in their selection of different linguistic forms as a means of symbolizing conceptualization (Hamawand 2016). Different frame facets that are specifically elicited as knowledge representation structures, in such processes, provide speakers with proper linguistic expressions for different construal (Hamawand 2016), thus dynamically

perform as temporary constraining factors for local meaning construction (Komatsu 1992; see also Johnson-Laird 1983; Barsalou 1987).

Compared with collocations, concordance lines provide more direct evidence of how *informational characterization* (Evans 2006) is realized by virtue of lexical concepts providing access to different kinds of knowledge structured in a particular frame. For example, an examination of the concordance lines of the word *generator* in ME sub corpus identifies three facets of the [GENERATOR] frame based on different sub-frames invoked by the participants for interpretive purposes:

- TYPES/FEATURES of generators (Figure 20);
- CHOOSING a generator (Figure 21); and
- EFFECTS upon generators (Figure 22)

Figure 18 TYPES/FEATURES of *generator(s)* (ME)

The first frame facet shown in Figure 20 is based on a shared descriptive knowledge of what a generator is, how it works and what features it has. The talk of TYPES of generators invokes the interpretive sub-frame of [ENERGY SOURCE] as the main criterion to differentiate a particular type (e.g. *an induction generator*, *a nitrogen generator*, *diesel generators*, *a megawatt generator*, *a tidal generator*). FEATURES and FUNCTIONS of generators show

the relevance and suitability to different situations, the talk of which invokes the [DISTINCTIVENESS] frame and the [TOOL PURPOSE] frame. Specifically, the relevance and suitability can be determined or influenced by a generator's STATUS (e.g. *main/auxiliary generators)*, SIZE (e.g. *bigger/larger/smaller generators*), SHAPE (e.g. *a cylindrical/square generator)*, SPECIAL PHYSICAL PROPERTY (e.g. *a waterproof generator*), PURPOSE (e.g. *a commercial generator*, *emergency generators*), CONDITION (e.g. *another new generator*) and DESIGN (e.g. *a conventional generator*).

The structure of knowledge activated by the above lexical concepts form the basis of an evaluative system concerning whether and why a particular generator is the best choice in a certain situation, which is also an important aspect in the participants' talk. See the concordance lines classified into the second frame facet: CHOOSING a generator to invoke the interpretive sub-frame of [CHOOSE] (Figure 21).

I'm just trying to think you do= so you don't really need a generator there? Eh. No because the generator will
to our uh vessel because we have also need to have a generator for a backup uh so as to achieve a better
I've designed a hub I've designed a shaft now I need a generator". Mm hm. Becu= but you design the shaft.
for really. Right. Right so would your if you picked out a generator could you actually say right this needs to
diameter eh I d= it may be worth if we can choose a generator or at least put say generator size speed that
is is rotating at a certain speed you then choose a generator accordingly. You don't try and change it. Well
be feeding that to Ellie who should then be choosing a generator from that and feeding it out to. Well we can
could take that torque value and choose an appropriate generator. Is that how it works? Technically I but you'd
Yeah well yeah we'll we will select obviously appropriate generators. Yeah. But like let's say on the program you
. The prob= the problem with that is we need Ellie's generator. Which is the thrust. Ellie can go after me
might from this research be able to pick like a suitable generator we could use for the tidal turbine as well. Well
the weight. I. All she needs to do is pick erm the generator gear box and the shaft sizes and then I can
you ch= you do the propeller first and then choose the generator not choose the generator and then. That's
would be erm electrical systems in helping to choose the generator the gear box stuff like that erm. But yeah it's
. Oh I thought you said you had picked. You picked the generator? No no I decided. Because my I spent this
blades. Well if you. The the rotor blades if you pick the generator you've basically picked the rotor blades cos
yeah. Yes. Yeah. And then from there you choose the generator. Yes I think so. That that's what we said
first and then choose the generator not choose the generator and then. That's what I think we should do.
generator from your propeller design not choose your generator first+ Yeah. +and then choose your propeller
that you you design your propeller first and choose your generator from your propeller design not choose your

Figure 19 NEEDING/CHOOSING a *generator* (ME)

The third frame facet is more complicated since it cannot be identified from the concordance lines that easily (Figure 22 shows some examples). This facet is about EFFECTS of various factors upon a generator, such as its working condition, service life span, maximized efficiency, etc. and the factors can be both internal and external. This facet concerns how the participants valued the [RELATION] frame (the interpretive sub-frame invoked) between the generator performance and other factors, which involved more cross-domain mappings for interpretation in local contexts.

Figure 20 EFFECTS upon *generator(s)* (ME)

This example shows that informational characterization has an interpretive consequence for natural language processing. The information characterization, however, is not equally rich or predictable for all lexical concepts. A high degree of specificity and predictability can be found with some lexical concepts (e.g. SHAPE, SIZE, STATUS, CONDITION, etc.) which provide access to a richly detailed semantic potential (Evans 2006: 509). By contrast, there are also cases in which a much lower degree of schematization can be retrieved from participants' prior experiences. Such experiences are thus linguistically represented by lexical concepts which may provide access that is more impoverished and more likely to emerge from context.

To summarize, this section provides evidence of the participants' invoking of specific interpretive frames to locally construct knowledge in their small group talk. The findings suggest that the interpretive frames invoked by the participants for situated meaning construction are partial reflection of larger-scale conceptual knowledge structures (Fillmore 1982a). Some of such structures are available and accessible from knowledge independent of the immediate communicative context. The semantic potential of lexical concepts provides access to the structures and determines what aspects of the structures are to be activated (Evans 2006; 2009) for local meaning negotiation. Prior to the immediate text such knowledge may be, the interpretive frames invoked are highly context-dependent and goal-oriented since they are discursively oriented to achieving certain goals. The goal-orientedness is a central characteristic of institutional talk (Drew and Heritage 1992; Heritage 1997; Heritage and Clayman 2010) and invoking proper interpretive frames is thus of particular importance to maintain the basic mode of interaction by making the interlocutors focus on the business at hand.

Another factor which constrains the participants from departing from the basic interactional mode is the fact that they have relatively equal access to the knowledge necessary to fulfil particular tasks. The epistemic statuses of the participants are found to be relatively stable upon a "flatter gradient" (Heritage 2012) of knowledge which includes their shared background knowledge about the task and their rights and responsibilities to claim the access to the knowledge. This ensures that the frame(s) invoked by individual participants are mutually interpretable for other group members and that meaning negotiation is optimized to achieve certain communicative goal(s).

This section has also shown that Corpus Linguistics methods are useful to approach patterns of lexical concept selection and interpretive frame invoking in natural language use. Collocations calculated on a frequency basis show great potential to reveal semantic tendencies of lexical concepts concerning the unique selectional requirements of their usage (Evans 2006). Concordance lines show the surrounding context of a target linguistic form in an organized way, which contributes to identifying patterns of the semantic preference of its associative lexical concepts.

5.4 Framing-in-interaction for knowledge construction

The research findings reported in Section 5.2 have shown that the identification of key semantic domains has an interpretive consequence for language processing in a sizable dataset like NUCASE. The over-used, non-shared key semantic domains are identified across the sub corpora through statistical tests and sample selections. In a most straightforward manner, they manifest the "aboutness" of the discourse in each sub corpus in terms of the knowledge which defines the scope and dimensions of the participants' small group talk. Linguistic forms calculated into the key semantic domains are examined to show what aspect(s) of meaning

is/are represented by the lexical concepts they encode. The relevant lexical concepts provide access to particular frames to characterize the discourse from the participants' points of view. This is the very first step the participants managed to make sense of the discourse: activating a relevant, coherent body of conceptual knowledge so as to transform their prior experiences into the basis of group collaboration.

The observations in Section 5.3 have revealed the relationship between frames evoked by linguistic forms and frames invoked for text interpretation (which is missing in Fillmore's discussion on cognitive frames). The analyses suggest that the interpretive frames invoked by the participants for situated meaning construction are the partial reflection of the frames evoked by the linguistic forms which are grouped within a specific epistemic domain. It is the semantic potential of lexical concepts that highlights the relevant meaning dimensions associated with the task goal, thus makes the larger-scale knowledge structures only partially activated and selectively managed.

This section illustrated how different patterns of meaning negotiation for knowledge construction emerge from the participants' applications of frames in interactional small group talk. A communicative dimension was added to the analysis of framing with different contextual factors taken into account. The analysis focused on sequences of the participants' talk around particular lexical concepts associated with linguistic forms which have high frequency values across the sub corpora. Section 5.4.1 examined sequences of alternate framings of a single situation. Section 5.4.2 dealt with framing sequences as a result of collaboration among the participants within/beyond a certain identity group. Section 5.4.3 focused on cases of framing in the participants' talk other than the alternate framings and co-framings. Interactional Linguistics analyses were conducted to approach the process of

knowledge construction in the moment-to-moment framing of the talk. The ways of framing are found to be not only associated with individual interpretations of the on-going discourse but also to manifest how a particular interpretation is made to reproduce a scene to one's interlocutors for mutual understanding.

5.4.1 *Alternate framings of a single situation*

The focus of this section is sequences which involve what is formulated by Fillmore (1982a) as "*alternate framings* of a single situation" (p. 386). The formulation is based on empirical semantic observations which show that the same 'facts' can be presented within different framings thus are made out as different 'facts' (Fillmore 1982a: 386). Alternate framings show how deviant people's experiential schematizations can be when they are encountered with the same situation. The single situation can make different people invoke different frames for situated interpretation. Fillmore (1982a) provided an example in which somebody who is unwilling to give out money in a particular context can be interpreted as *stingy* in contrast with *generous*, or as *thrifty* in contrast with *wasteful* (p. 386). The former interpretation evaluates the behaviour in terms of its effects on others by invoking the [TREATMENT] frame; whereas the latter evaluates the behaviour in terms of the wisdom of using money by invoking the [FRUGALITY] frame. The two ways of evaluating the same behaviour are both partial representations (Coulson 1997) of how a social encounter enters into individual conceptualizations.

Excerpt 1 provides a good starting point to examine framing in such way. The alternate framings in this example, however, may not be taken as typical since the two different ways of interpreting a single situation are provided by a single speaker rather than proposed by the interlocutors. In this excerpt, a tutor ($1) and two student-teachers ($5 and $7) are talking

about student-teachers in the workplace in a PGC teacher training feedback session. The aim of the session is to elicit from the student-teachers a reflection of how they have been treated by senior teachers in the workplace and how the treatment adds an emotional dimension to their career experience.

Excerpt 1. ***student-teachers in the workplace***

```
1   <$1>  So e= I ↑guess (.)in a way(0.5)for a= for a= se:nior
2         teacher to come into your lesson >as you say first of all
3         you think< "oh my word↓ that's (0.6) you know (.) bit
4         [↑worrying"]
5   <$7>  [Yeah      ]
6   <$1>  (.)but then ↑actually >the fact that< he gave: you: the
7         ↑respect to say ["well      ]
8   <$5>                  [Mm (.) Yeah]
9   <$1>  what= what <do you:: want me to [do?"]>
10  <$5>                                  [Yeah]
11  <$7>                                  [Mhm ]
12  <$1>  =given that we've (.)There's obviously >an incident<
13        going on here =There's ↑always gonna be incidents(1.0)
14        but rather than just saying "oh you ↑°silly°
15        [student teacher]
16  <$5>  [Yeah           ]
17  <$7>  [Yeah           ]
18  <$1>  =and >you know< I'll sort ↑this out for you"
```

The discussion is based on a shared acknowledgment among the participants that there is a tension between student-teachers at work and other institutional actors (e.g. senior teachers) concerning how the former is viewed and treated by the latter. The single situation the participants are dealing with is "a student-teacher's lesson is under observation of a senior teacher". The [EXPECTATION] frame is evoked as contextually relevant concerning predictions from the both sides of "what is going to happen" in such a lesson. $1 first provides one way of interpreting the situation from the student-teachers' points of view, namely, that

the student-teachers would feel unconfident and worried to teach before an expert. The [WORRY] frame invoked by the tutor is confirmed by $7 (*yeah*, Line 5). $1 then immediately proposes the other way of interpreting the situation by invoking the [RESPECT] frame to show that "being observed by a senior teacher" can be something positive since the student-teachers are in fact helped rather than criticized. The word *silly* (Line 14) stages an external voice which has been pragmatically revised with a rising intonation but in a lower volume. It performs as a counterfactual marker, indicating that the situation being discussed at hand is contrasted.

Sensitivity to the institutional identity conflict is manifested in this excerpt. $1 as a tutor tries to ease the tension between student-teachers and senior teachers by portraying herself as an "on-looker" instead of "one of" the latter, aiming to express a relatively neutral position. This can be seen as a positive politeness device (Brown and Levinson 1987) to weaken the inclusiveness of herself in the senior teacher group. Despite the fact that she is entitled to speak authoritatively with a tutor membership (Potter 1996), she deliberately chooses to suggest "closeness and equality rather than distance and hierarchy" (Benwell and Stokoe 2002: 436). However, interestingly, $1 seems to imply that such "respect" may be something that is simply performed as a face-saving act. The rhetorically expressive talk can be found in her change of intonation and stress of the word *respect* earlier (Line 6). That senior teachers try to be considerate and nice to student-teachers is thus perceived as something for politeness. This, in turn, may suggest the power bias stabilized as an institutional practice in the disguise of mutual understanding and respect.

In Excerpt 2, one shall see how alternate framings are applied from different speakers who share a particular identity which is institutionally defined. In the excerpt, two tutors ($1 and

$2) and a student-teacher ($4) are reflecting upon a scenario recalled by the student-teacher from her prior teaching experience. An orientation of $4 can be identified to build up her professional value of being a teacher with students' emotional feedback.

Excerpt 2. ***the teacher-student emotional bond***

```
 1  <$4>  I ↑want to get to the stage where the kids they're sad(.)that
 2        I'm not gonna be teaching them °anymore°((laughing))
 3  <$2>  [They >probably< sure they were sad when you left(.)
 4  <$1>  [I'm sure they al= they al= they [↑ALREADY WERE         ]
 5  <$4>                                   [↑Yes some of them WERE]
 6        (0.5)I was really pleased(0.5)and I was like= "ah yeah(.)my
 7        lesson must be quite good (.)if they [think= if they really
 8        said that"
 9  <$2>                                       [Yeah(.) But-
10  <$1>                                       [Yeah(.) Well=
11        that's what Roger and ↑I ho:pe (.) ↑isn't it(.) >That at the
12        end of the year< you think-"aw:I'm gonna miss them"
```

It is right after $4's turn where alternate framings carried out by two tutors can be identified. The overlapping talk (Line 3-4) shows their immediate responses to $4's utterance. $2 talks about the POSSIBILITY (*probably*, Line 3) of the kids being sad on the kids' part, while $1 emphasizes the CERTAINTY (*sure*, *already*, Line 4) of the kids being sad from her own point of view. $2 invokes the [EMOTION_DIRECTED] frame to objectively examine the bond of the teacher-student relationship; whereas $1 invokes the [EXPERIENCE_FOCUSED_EMOTION] frame to show her empathy with $4. $1's comment seems more proactive and encouraging, indicated by $4's acknowledgement (Line 5) which relates the kids' reaction to positive self-evaluation (Line 6-8). Note that $2 orients to maintaining his framing by trying to give a different comment after a short acknowledgement (*yeah*, Line 9). The act is projected by the word *but* (Line 9) which indicates that $2's following talk will be contrasting with his acknowledgement. This move, however, is interrupted by $1 when she may have realized that what $2 is going to say would probably

discourage $4. She then deliberately applies a series of discursive devices to "save the talk": including $2 into the same group with herself (*that's what Roger and I hope*, Line 11), using a tag question (*isn't it?*, Line 11) to check confirmation (see Gass et al. 2005) from $2, and directing the talk into a sympathetic realm (*Aw, I'm gonna miss them*, Line 11).

The overlapping talk in Line 9 and 10 shows individual efforts made to maintain their framings which have been constructed in the prior talk: $ 1 orients to maintaining her framing with $4 while $2 orients to regaining the focus on his framing but fails to do so when $1 manages to take the floor. The observation suggests that alternate framings can be pre-staged by the use of particular discourse makers at turn initial positions. In this excerpt, specifically, the acknowledgement token *yeah*s at the initial positions of both $1's and $2's turns are associated with a display of passive recipiency which exhibits "a preparedness to shift from recipiency to speakership" (Jefferson 1983). The possible next actions projected by the same token *yeah*, however, can be contrastive: it could be a positive evaluation ($1) or a shift of focus ($2). This is the structural representation of the cognitive divergence involved in alternate framings of a single situation.

Excerpt 3 is an example to show that alternate framings of a single situation can be associated with different mind sets shaped by distinct institutional values. In this excerpt, a student research team ($2 and $4) and a company delegation ($1 and $3) are talking about their concerns of a software design that the former is expected to submit to the latter. The focus of the talk is the structure and content of the design documentation.

Excerpt 3. ***software development cycle***

<$1> ↑Well(.) so another requirements gathering meeting(.) Is there anything(.) er(.) you need to know ahead of your(.) submission(.) of the initial(.) design this evening

<$2> We're just looking for clarification on what it was(.) with regards to the documentation that you actually wanted(.) Like- obviously the aims of each(.) er(.) proposal and also(.) which tools we're going to use ((a female's coughing))and why we're going to use them(.) Is that mainly what you're ↑after

<$3> (4.0) Design documentation and er: I guess(1.0) have you looked it up on ↑Google(2.3) The software development life cycle(.) Do you know what design documentation will look like(.) I think do you ↑KNOW(.) Well I hope you do because I'm paying you enough(.) about the software development life cycle(.) You have one at least= one computer scientist on your team

<$2> Yeah(2.3) Okay-

<$1> =So there are standards(.) for design documentation(.) ((a male's coughing))and I think what we'd like to see(.) is(0.5) ↓documentation that conforms to those standards

<$2> Okay(4.8) >That was the only question I really came in with<= I ↑mean(.) the rest of it is just getting on(.) with the work flow= so-

<$1> Right

<$4> Yeah(.) it's pretty much(.) the same(.) as yesterday we found

<$M> Mm-mm

<$3> So what a= a company would be looking for is some evidence that there is some rationale from(.) because we all put our contracts-

The two parties are found to frame differently a single situation of "a software design is to be presented in documentation". The research team tends to highlight in the documentation the contextualized factors (e.g. aims, tools, rationale) which were taken specifically relevant and significant to their design. On the other hand, the company delegation prefers a reference to a standard model of software development life cycle which will specify and rationalize the

order of stages of the software design. The research team frames the process of producing the design documentation as a BOTTOM-UP one while the company delegation frames it as a TOP-DOWN one. Conceptualization of "expertise" is represented from different perspectives concerning what should be the common practice in the field. The company delegation claims the existence of a knowledge gap in the research team and highlights the responsibility of the latter as employees (Line 10-16).

Evidence can be found that the both sides may be reluctant to accept each other's ideas of how the documentation should be like. This is revealed by three remarkably long gaps. One of them emerges in the *transition space* (Sacks et al. 1974; Jefferson 1986; Schegloff 1996a, 1996b) for speaker change (Line 10). The other two emerge within the research team's turns (Line 17, 21). The three gaps, however, show different ways the talk is sequentially manipulated. For the first gap, the company delegation expands the transition space by not providing talk which has been projected by prior talk (Liddicoat 2007). This is interactionally relevant and indicates that they do not really agree with the research team. An alternate framing is then provided by the company delegation right after the gap. The second gap after $2's acknowledgement (*yeah*, Line 17), along with another acknowledgment token (Line 17, *okay*), is perceived by $1 as a transition relevance place where any participant can legitimately take the floor. The silence here is attributive to $2's not speaking and showing his failure to maintain his original framing. The company delegation, on the other hand, manages to stay in the frame they have applied to the talk earlier (line 18-20; 28-30). The third gap which is also after $2's acknowledgement (Line 21) interactionally creates another prolonged transition space for possible speaker change. The transition space, however, develops into an "intra-turn silence" (Liddicoat 2007: 81) with $2 tending to abandon his original framing,

which is marked by his use of *okay* (Line 21) to propose a readiness to move out of the current topic (Beach 1993).

Institutional values are relevant to the analysis of this excerpt. The alternate framings show the different views of the degree to which a design work should conform to well-established standards. The research team is not familiar with commercial popularization and application, while the company delegation is not familiar with inquiry principles and methods of scientific research. The knowledge gap is originated from and may be expanded by institutional barriers, at both the conceptual and practical levels. Such perspectival conflicts are found to occur in other institutional discourses (e.g. police interrogations, Jönsson and Linell 1991; court trials, Gumperz 1995) where participants seem to have different tacitly held contracts about their language behaviors and discursive predictions (Linell 1998b).

We also have cases in which alternate framings of a single situation simultaneously emerging from interactional sequences are the result of a conceptual mismatch between the prior speaker and the current speaker. The conceptual mismatch can be explained by the prior speaker's particular lexical choices at the point where speaker change is relevant, which makes the current speaker think that the following talk is projected into a path for a contrastive interpretation. In Excerpt 4, two students are talking about the consequence of an over-compensated generator in a monitoring system.

Excerpt 4. ***the generator getting too hot***

```
1   <$4>  =So do we almost have to O:VER-(1.6) °what's the ↑word°
2   <$2>  Overcom=
3   <$4>  ↑YEAH(.) Overcompen↑sate
4   <$2>  Well the o= the only problem with that is(0.5) um(0.5)
5         obviously at rated torque is the most efficient(.)You
6         know at the what- the rating of the generator determines
7         the(most)efficient-
8   <$4>  =Right so you want it working at its pea:k
9   <$2>  You w= you want it working at its peak(.) but obviously-
10  <$4>  [You've got a heat problem
11  <$2>  [=you have cooling systems-
12  <$4>  =Yeah(.) ↑AH ↓RIGHT(.) Okay
13  <$2>  You ↑know
14  <$4>  Mm
```

Prior to the beginning of this excerpt, the [TEMPERATURE] frame was evoked when the central topic of the project meeting was introduced to the participants, namely, the temperature of a generator to be monitored. This frame highlights a certain externally verifiable state where the entity is observed and described, rather than its experience of the temperature (see FrameNet Project). Such an externally verifiable state defined by temperature is closely associated with the generator's performance at work. Note that although the temperature state is externally verifiable, it is the internal mechanism of the generator that determines how the state is brought into being at a particular time, how it can be maintained or changed in some way and how the maintenance or change will have an impact.

The talk then progresses around a single situation: "the generator is getting too hot" with $2 and $4 collaboratively retrieving the term *overcompensate* (Line 3) from their shared prior knowledge to describe the situation. In Line 4, $2 suggests that it can be a problem to keep the temperature low at one hand and to have the generator to work most efficiently at the other

hand (Line 5). A temporary "meeting-of-minds" (Gärdenfors 2014: 77) is achieved when $2 confirms $4's interpretation that the generator is expected to be working at its peak (Line 8, 9). The linguistic form *but* uttered by $2 at the end of his turn (Line 9) pragmatically operates to display a possible action completion for 'contrasting' what has been already constructed in his prior talk (Hata 2016: 139). The contrast sequentially projects a stepwise move from $4's point of view, encouraging him to go back to the situation of "the generator's getting too hot" in their earlier talk. Following the hint, $4 invokes a [CONTRAST] frame by focusing on the contrast between the preferred working status of a generator (*working at its peak*, Line 8) and its dis-preferred consequence (*You've got a heat problem*, Line 10). $2, however, alternately frames the talk by invoking another frame of [CONTRAST'], focusing on the contrast between the problem and the solution (*You have a cooling system*, Line 11). $4's follow-up turn with his saying *Ah right* (Line 59) as a reception marker (Smith and Jucker 1998; Fuller 2003) indicates that he is informed (Hata 2016; see also Schiffrin 1987) of what $2 really means, marking a change in the "locally current state of knowledge of awareness" (Drew and Heritage 1992: 41). The transaction of information is made relevant and mutual understanding adequate at this point where the minds meet with an attendance to the ongoing interactional concerns among the interlocutors (McCarthy 2003).

The observation suggests that alternate framings of a single situation can be the result of sequential manipulation of interaction at the action level. $2 and $4's overlapping talk reveals how information is gathered, interpreted and conveyed from different viewpoints (Minsky 1977) even when they have similar pragmatic orientations (e.g. to express contrast). The overlapping talk cannot be simply taken as something that $4 entering the talk does to $2 who currently has the floor thus makes the interaction problematic. On the contrary, the overlapping talk as an interactional phenomenon has an interpretive consequence for alternate

framings around a particular situation emerging from the progressive talk. The analysis also manifests that negotiation of epistemic stance is subject to the moment-to-moment expression of institutional relationships for situated knowledge construction (Heritage 2012; 2013).

Finally, let us have a look at one of the cases in which alternate framings of a single situation can be complicated with the interlocutors managing the mutability of their framings. In Excerpt 5, an expert in biology ($7) employed by a pharmaceutical company and a student research team ($7, $8 and $9) are talking about using public databases for drug re-purposing. One shall find in this example that representations of professional identity are shaped and constrained by institutional factors and that the common interest embedded in the profession is expressed in different ways.

Excerpt 5. ***using public databases for drug re-purposing***

<$8> You see the thing is- I guess- that these(0.8)these online databases are= are all public knowledge so-
<$7> That's right
<$8> =it wouldn't be an e= an ex= any exclusive rights for you(.) For yourselves= Other than the fact that you've already got the drugs developed
<$7> Er thi-
<$8> =Er unless- the er- no unless >°the drugs are sort of in the public domain already°< and there's no(.) patent for them
<$7> ↑No this is true(.) It's a good ↓point(.)but erm(.) It ↑is a good point ((laughter))
<$9> I guess= I guess being first in mar:ket actually
<$7> Yes
<$9> It makes a huge difference(.)trying to market them

((23 lines omitted))

<$7> Ev= every but there's SO MUCH information out there in databases(1.0) that <no: one: has got the same view:> of what's in there= What ha= tends to happen is people do these high ((unclear speech)) experiments(.) They put their data in the databases(.) Quite often it never makes it into the ↑literature(.) So as bio info -↑medicists(.)we have this huge chance(.) of pulling out new ↑knowledge from existing data(.) So the data exists(.) but the knowledge about what it might ↓mean(.) in a biological context(.) doesn't necessarily(.) And ↑this is what we're trying to do ↓here
<$8> (2.0)Okay

In the talk prior to the excerpt, $7 as the biological expert explained to the research team the rationale and advantage of drug re-purposing on the company's part; namely, re-purposing a drug which has already been approved for one purpose can save a lot more time and money than getting a new drug approved by law. This means that getting existing drugs re-purposed can help a company make profits at a lower cost in a shorter cycle of research and

development (R&D). The talk then progresses around a single situation: “the company does not own the copyright on the knowledge in public databases”. The financial dimension highlighted by $7 in her earlier talk is perceived by the research team as relevant to the project they have been undertaking. They evaluate the situation under focus as something negative for the company’s research endeavors, thus frame the situation by highlighting the necessity of prompt patenting and marketing of the drugs which are already developed so that the company can take advantage of exclusive rights to the data for further research on repurposing. The framing practice associates *marketing* with the financial concerns of the company. $7 as an employee of the company, by contrast, alternately frames the situation as something positive concerning its potential to benefit medical research in the long term. She emphasizes the possibility of *expanding professional knowledge* through an active and creative exploitation of public databases. In such sense, the inclusive rights for all social actors make the knowledge in the databases accessible to a larger scientific community thus can promote academic communication and economic growth.

The alternate framings in this excerpt show the participants’ orientations to departing from their identities which are institutionally defined. What one would probably assume is that the student research team would pay more attention to the expansion of academic knowledge while the company delegate would be more interested in making profits. The observation above, however, shows the opposite. We have seen that the both sides in the talk attend to each other’s conceptual needs and make the alternate framings mutually accessible and acceptable. This can be partially explained by the fact that the student research team is bound by the project funded by the company and that the company delegate shares the core values and principles in the professional field.

5.4.2 *Co-framings within/beyond the role boundary*

In this section, the focus is on how the participants in small groups collaboratively framed the talk to make it progress. I shall call such kind of framing *co-framing* which is motivated by a shared orientation among participants to a particular task. Different from alternate framings discussed in the previous section, co-framing shows a closer association with speaker roles which are either assigned in the task script or naturally emerging throughout interactional sequences (see Dörnyei and Murphey 2003). A role implies the relationship between one's actual behaviour and the shared expectations from other relevant members in a group. In the small group talk scenarios I shall be looking at, the speaker roles are either assigned within an institution (e.g. tutors vs. students), across institutions (e.g. research students vs. company delegations), or by task specifics (e.g. different roles assigned to students to complete a project). The role assignment ensures that each of the participants in a group has got "something specific to do" (Cohen 1994) to increase learning potential of the group which is essential for task completion. On the other hand, the establishment of emerging roles is a powerful component of group interaction (Forsyth 1999) which can reveal the contextual relevance of co-framings both within and beyond the role boundary.

I shall start by dealing with one instance in which co-framings are carried out by participants with a specific task orientation. In Excerpt 6, two participants are talking about how to draft their project report on the part of calculation of wave loading. Prior to this excerpt, they talked about the inquiry focus and the division of team work.

Excerpt 6. ***reporting wave loading calculation***

```
1   <$2>  =Well for me- well- the way when I pitch it if I have to
2         talk about the stuff I ↓do(.) I will tell them(.) what I
3         ↑had
4   <$1>  Yeah
5   <$2>  What I had to deve↑lop
6   <$1>  Yeah
7   <$2>  To work out the results
8   <$1>  Yes
9   <$2>  What I did have is the= a class report
10  <$1>  Yeah
11  <$2>  With the: significant wave height
12  <$1>  Yeah
13  <$2>  Then I ha:d to look for ↑formulas-
14  <$1>  Yeah
15  <$2>  =to find the wave ↑length-
16  <$1>  Yeah
17  <$2>  =and the wave- eh(.) whatever characteristics of the ↑wave-
18  <$1>  Yeah(.) yeah
19  <$2>  =and use Morison's equation-
20  <$1>  Yeah
21  <$2>  =to develop the= the= the wave ↓loading
22  <$1>  Yeah
23  <$2>  The current loading the wind loading u= works on about the
24        same ↑principle-
25  <$1>  Yeah
26  <$2>  =used the Atlas and so on and so forth
27  <$1>  Yeah
```

The two participants then pay a joint attention (Gärdenfors 2014) to reproducing the procedure of the report by highlighting the transactional[22] dimension of the report. It is evident that \$2 is the person who is responsible for reporting the calculation of wave loading. As his partner working on the same part of the project, \$1 is expected to facilitate \$2's reporting by simultaneously monitoring the process to check the accuracy of information and the logic of inquiry. The roles assigned to the task results in the following linear sequences in their talk, through which \$1's responses to \$2's utterances are of particular interactional relevance. The turn-takings are quite rapid and compact with \$1 using the response token *yeah* (see Gardner 2007) 12 times (*yes* for once, Line 8). Schegloff (1982) observed the

[22] Transactional discourse is typically characterized as using specific lexical items to provide 'an optimally efficient transmission of information' in written and spoken interaction (Brown and Yule 1983; Lakoff 1989).

multifunctioning of the response token *yeah*: it not only marks acknowledgment and confirmation but also expresses agreement. On the other hand, the repetitive use of *yeah* in $1's responses over $2's extended talk, as suggested by Schegloff (1982), could run the risk of being taken as a sign of the hearer's feeling bored or losing attention. However, McCarthy (2003) suggests that *yeah* used as repeated token could be interpreted as "signalling an enthusiastic or encouraging response" (p. 40).

A reasonable interpretation of the repetitive use of *yeah* in this excerpt, however, requires analysis of the token along with other contextual resources to explicate its affective consequences (McCarthy 2003). According to Gardner (2007), *yeah* as a response token is "more retrospective than prospective" and reveals more involvement and more speakership incipiency (Jerferson 1984). I shall argue, in this case, that $1's repeated uses of *yeah* indicate a combination of both affective attendance and communicative economy (Schegloff 1982). $1 uses the *yeah*s as continuers (Schegloff 1982; Goodwin 1986) to construct his concurrent talk (Goodwin 2007), frequently informative about his analysis of what is being said by $2 and his stance towards it (Jefferson 1983; 1984).

On the other hand, $1's right to take turns is to a large extent constrained anyway (see Houtkoop and Mazeland 1985; Schegloff 1982) when $2 is engaged in an extended report which is tightly bound by the task procedure. $1 says *yeah* repeatedly, but within "quick and close sequences" (Tottie 1991: 261), which indicates that encouraging $2 to go on talking is possibly due to the consideration of communicative economy. In other words, $1 intends to make the discussion as concise and efficient as possible by holding $2 back from further extending his turns. In sum, $2's framing is information-loaded while $1's framing is

procedure-loaded, both of which orient to completing the task at hand, namely, rehearsing for a formal report.

Now I turn to examine cases in which co-framings are applied by participants who are actors within a single institution. Excerpt 7 shows how two tutors ($1 and $2) are collaboratively explaining what schools expect of student-teachers. Prior to the excerpt, the tutors and the student-teachers talked about the fact that senior teachers who have been in the career for a long time may have almost lost their passion and sensitivity to students' feelings.

Excerpt 7. *what to expect of student-teachers*

1	<$1>	↑Well and also I think >a lot of them< ↑appreciate how much
2		tou:gher it is(.) You know(.) what the expectations are on
3		student-teachers
4	<$5>	Mm(.) Yeah
5	<$1>	(1.5)Um and I we= I think ↑we forget (.)actually(.) about-
6		[what= what= <the SYSTEM now EXPECTS of you>]
7	<$2>	[°The pressure in schools has increased°](.)Yeah
8	<$1>	(.)I mean- when even three ↑years ago(.) we >didn't have to
9		use< the Ofsted um(.)[criteria for-
10	<$2>	[That's right(.)Mm
11	<$1>	=satisfactory good and outstanding at student-teacher
12		level

At the beginning of this excerpt, $1 initiates the topic of institutional expectation of student-teachers by arguing that student-teachers nowadays are met with more challenges (*how much tougher it is*, Line 2) since nearly all schools have begun to adopt standard systems to evaluate their professional capability and teaching performance. The mounting challenges, as claimed by $1, are expected to lead to increased pressure felt by student-teachers when they work in the school. The consequent stress is identified by $2 whose utterance (Line 7) can be

seen as agreement with $1's argument. $1's overlapping talk (Line 5), however, indicates that what she is concerned about is the gap between the existing evaluative systems and what to expect of student-teachers in reality. The overlapping talk may be perceived by $2 as something problematic when he realizes that $1 and himself would probably push the following talk into different conceptual realms. He chooses to close his turn after a short pause with an agreement token *yeah* (Line 7). $1 then takes the floor to build upon her prior talk by making it clearer (*I mean*, Line 8), pointing out that the explicitly laid-out criteria in a standard evaluative system (e.g. the Ofsted criteria[23]) may not be more useful or reliable than what schools did before the system was introduced. Her idea receives a confirmation from $2 (*That's right*, Line 10) which is uttered in an overlapping way again. The shortened transition space here, however, can be seen as attributive to $1's short pause (Line 9) which seems to create a place for legitimate speaker change. Finding that $1 orients to holding the floor after her pause, $2 again chooses to close his turn to make the talk progress.

The observation reveals that at a particular moment of an interaction a leading role (e.g. $1) may naturally emerge to frame the talk while other relevant participants (e.g. $2) can choose to challenge or follow the leading role by manipulating the talk sequences. In this example, $2's co-framing practice with $1 is represented by his following and building upon the latter's talk, even though the contextual relevance of taking over her leading role is made semantically and pragmatically salient to him. While $1 tends to produce extended turns within her frame, $2 manages to make his turns short and brief to maintain the progressivity of the talk (Brandt 2011). This is how co-framing is carried out not only at the level of meaning but also at the level of action.

[23]Ofsted is the Office for Standards in Education, Children's Services and Skills. We inspect and regulate services that care for children and young people, and services providing education and skills for learners of all ages (Ofsted official website https://www.gov.uk/government/organisations/ofsted/about).

Excerpt 8 is another example to show how co-framing can be applied at the action level but beyond the institutionally specified role boundary. In this excerpt, an expert in biology ($7) and a member of a student research team ($8) are talking about what to find in drug targeting. In the prior talk, $7 has explained to the research team the similarities and differences between a drug target and a vaccine target and highlighted the fundamental logic for carrying out the project.

Excerpt 8. *what to find in drug targeting*

1	<$8>	Yeah(.) But that's why we in the first one we're looking for
2		variants and this one we're just looking for(.) erm(.)
3		alignments-
4	<$7>	Right
5	<$8>	=To see(.) what level of alignment we've ↑got(.) Erm(.) but
6		also when I- said characterise in the first one= We was
7		characterising for(.) basically location and accessibility-
8	<$7>	Mm-mm(.) Sounds good
9	<$8>	=And this one(.) we're looking for metabolic function
10	<$M>	Mm-mm
11	<$7>	Right
12	<$8>	So-
13	<$7>	=Nope(.) That sounds like a= a reasonable approach(.) ↑Yeah

$8 as a member of the student research team then proposes their approach to the two parts of the project based on the research logic explicated by $7. $8's extended turns (Line 1-3; 5-7; 9) receive brief acknowledgment and comments from $7 (*Right*, Line 4; *Mm-mm. Sounds good*, Line 8; *Right*, Line 11). Note that the two *right*s used by $7 as response tokens can be understood as *epistemic dependenc*y markers which reveals her recognition of the relationship between what is currently under discussion and something that had been said earlier (Gardner 2007: 325). It is in Line 12 that $8 orients to extending his talk by initiating the turn with *so*

(Line 12). Because his turn is interrupted by $7 right after his utterance of *so*, his following action is open to multiple predictions. He may orient to introducing the result of approaching the project in the way he has just mentioned; to clarifying the motivation for adopting the proposed approach, or to providing an evaluation on its rationale. His framing is constrained from further expansion with $7 entering the interaction.

The word *nope* (Line 13) is quite curious considering what $7 says following the word: *That sounds like a reasonable approach* (Line 13). A conceptual conflict can be identified that she negates what the last speaker said and shows an affirmative attitude right after the negation. The hidden psychological process becomes traceable and interpretable when one goes back to examine the prior interactional sequences. One possible interpretation is that what $7 negates is not what $8 said but her next move to give comments. The fact that $8's talk is interrupted indicates that he prefers another extended turn over comments from $7, while $7 might be ready to comment from the moment $8 began his talk in Line 1 but decides not to do so. This could partially explain why $7's replies since Line 1 are quite brief – she has been considering $8's proposal and does not want to suspend her train of thought by stopping to give longer comments. This is also revealed by $7's use of the word *nope* but not the less emphatic *no*, by which she has no intention to change the truth condition (see Schourup 1999; Fuller 2003) of $8's talk but still uses the negation mark to show her agreement in a relatively relaxed manner. In addition, *yeah* with a rising intonation at the end of $7's turn (Line 13) suggests that an acknowledging action in response to the prior other's action is embedded as a cognitive consequence.

The observation shows that a conjunctional (e.g. *so*) may be vulnerable to another speaker's turn initiation (Jefferson 1983) since its semantic potential can be pragmatically rich thus

leads to multiple interpretations of what is going to happen next. In such situation, it is possible that the next speaker would see it as a legitimate place for speaker change and reduce the transition space accordingly to express his/her interpretation. Co-framing can, as a result, be challenged if the next speaker's interpretation happens to be divergent from what the current speaker means. This would have been the case in this excerpt if $7's follow-up agreement was missing since a reduced transition space and a salient negation mark (*nope*) are commonly seen in cases of disagreement or rejection of prior talk (Liddicoat 2007: 86). However, we cannot claim that the co-framing in this excerpt is successfully achieved but that one possible way of co-framing has been applied at the action level. $7's utterance does have conceptual relevance to $8's prior talk and to the whole sequential organization since it projects the equal probability of $7's next action which could either be minimal responses or an extended turn of comments.

The phenomenon is related to what is coined by Schegloff (1972; 1992) as *conditional relevance* in sequential organization of talk-in-interaction. Schegloff (1992) described the notion of conditional relevance as "a first action creates a slot for an appropriate next action such that even the absence of that action can be perceived as a relevant and noticeable event" (p. 191). In such sense, whether or not the next action occurs (and in what way it occurs) can be seen as an event not only with "noticeability" but also "legitimate and recognizable grounds for a set of inferences" (Schegloff 1972: 76). If a next action is appropriately taken as responsive to the first action in a certain sequential organization, a relational preference (Atkinson and Heritage 1984) projection is expected to be established within adjacent turns. According to Liddicoat (2007), such preference projection has nothing to do with interlocutors' personal desires for meaning transfer but relates to the "recurrent patterns of talk in which actions are carried out" (p. 110-111). Some actions in a particular context can be

encouraged while others are to be avoided. Why and how certain actions occur at a certain stage of interaction is closely related to the routinely performed speaker roles which are institutionally specified. In Excerpt 8, for example, the negotiation of epistemic stance (Heritage 2012) shows reverence for the knowledge authority ($7 as an expert) and the more powerful social group (the company as an employer) which are associated with symbolic role assignments (Brewer and Gardner 1996).

We also have cases in which co-framing is conducted when individual framing moves are not taken at the same level but in a hierarchical structure. For example, Participant A first invokes a particular interpretive frame at a particular moment of an interaction, and then Participant B co-frames the talk by invoking another frame which is embedded in Participant A's frame. When participant A changes one or more elements of his/her frame[24], participant B adjusts his/her interpretation accordingly. Excerpt 9 provides an example. In this excerpt, two participants are talking about monitoring the temperature of generators at work. Their goal is to prevent a generator from destructively high temperature but at the same time to obtain its maximum efficiency.

[24] A frame element refers to a frame-specific defined semantic role that is the basic unit of a frame (FrameNet Glossary).

Excerpt 9. ***temperature monitoring***

```
1   <$4>  I was just thinking about tidal turbines= How they're
2         probably gonna be constantly moving(.) Can a generator
3         get to a certain temperature and just stop wor:king(.) Or
4         w= will it ha= it'll have to be shut down before it gets-
5   <$2>  =Yeah it'll have to be= er it'll break(.) otherwise
6   <$4>  Yeah
7   <$2>  Because um basically a g= a generator is rated(1.0)
8         and(.) at a certain ↓torque(1.0) And that torque is
9         determined by(.) the tempera↑ture(.) So you could run at
10        a higher= higher than rated torque-
11  <$4>  Right
12  <$2>  =for let's say five minutes(.) as long as you've been
13        running(.) twenty minutes below rated torque
14  <$4>  Right(.) Okay
15  <$2>  So it- you will have to obviously select a t= er a
16        turbine(.) that will let you do (what) generator that
17        will let you do ↑that
18  <$4>  ↑Yeah
19  <$2>  Um(.) but they will be(.) on mo= monitoring systems as
20        well(.) So as soon as this temperature excee:ds what(.) X
21        temperature(.) it will then have to be shut down(.) OR:
22        just run at less than rated
23  <$4>  I mean(.) if it's producing it(.) um: it's (2.2) if it's
24        producing its rated ↓tor:que
25  <$2>  Mhm
26  <$4>  Um will(1.8) will it just get to a certain temperature(.)
27        and stay (there) or will(1.8) continuous(1.5) use of it(.)
28        eventually get the temperature too ↑high
29  <$2>  Yeah
30  <$4>  It ↓will(.)become too ↓high
31  <$2>  ↑Yep
32  <$4>  Right
33  <$2>  Eventually
```

The selection of tidal turbines is taken by $4 as central to achieve their goal and the working mode of tidal turbines is closely associated with the generator design. $4 thus frames his talk by focusing on the design features of a generator, with two relevant frame elements presented in contrasting ways; namely, when a generator gets to a certain (high) temperature it will a) stop working by itself, or b) needs to be shut down through external intervention. $2 confirms the latter situation by latching his talk to $4's talk. The rationale of imposing intervention to

shut the generator down is thus made salient to interpret the following talk. $2 then invokes two sub frames associated with the "design features of a generator" to explain why it is the case (*because*, Line 7) and how the case can possibly apply to turbine selection (*let's say*, Line 12). Specifically, as suggested by $2, a generator can run at a torque higher than the rated torque in its speed design (Line 7-10) and the temperature of a generator can constantly change according to the changing torque created by the tidal turbine. In Line 19, $2 re-confirms the second element in $4's framing as one of the consequential responses to the monitoring system and completes his co-framing through a partial reflection on $4's prior framing.

However, $4 in Line 23 projects the interaction back to $2's prior talk about the rated torque of a generator. This move indicates that $4 may have identified from his original framing something associated with the concept of *rated torque* as a major source of trouble which has made the framing not readily accessible to $2. This is proved by $4's initiating a repair utterance (*I mean*, Line 23) (Schegloff 1997) to try to re-formulate his original framing. This repair move is rather delayed because it is after $2 expanded his turns that $4 has located the trouble source. Nevertheless, it shows a certain sequential connection between the current talk and the non-adjacent yet "not-so-distant" prior talk (Sacks 1992: 349-350). In his following talk (Line 26-28), $4 re-formulates one of his original frame elements "(a generator will) *just stop working*" into "(it will) *stay there*". The new frame element and the original one are pragmatically contrasting with each other, which shifts the interpretive focus of $2 from "how a generator will be stopped" to "whether a generator can self-control temperature". The re-formulation triggers $2's re-interpretation of $4's framing and his adjusted co-framing move (*yep*, Line 31) invited by $4 (*It will become too high*, Line 30).

The observation in Excerpt 9 shows that the process of meaning-making for knowledge construction can be iterative rather than linear at the sequential level and that co-framing among participants may involve moment-to-moment re-formulation and re-interpretation of individual framings in the interaction. Specifically, the reformation of framing can be realized using "explicative contrasts" to "warrant a deviation categorization" (Deppermann 2005: 301). Such contrasts are not necessarily achieved based on the context-free semantic relations between lexical units but are pragmatically constructed for situated meaning-making. In this excerpt, $4's original frame element *a generator will stop workin*g is violated by the re-formulated frame element *a generator will stay there* with the latter functioning as "a semantic clarification of the local meaning" (Deppermann 2005: 302). The two linguistic forms *stop* and *stay* cannot be readily mapped onto a specific semantic dimension to represent the meaning of contrast. They are, however, co-selected in this context to instruct the hearer ($2) to adjust his interpretation and, accordingly, his following action. The pragmatic opposition within the frame is built up in the interactional sequences and facilitates discursive co-framing for local knowledge construction.

As discussed in Section 5.3.2, some lexical concepts can serve to encapsulate "often complex and informationally diffuse ideas" (Evans 2006: 509). This is one dimension of their semantic value called *encapsulation*. Co-framing to make sense of such lexical concepts can be challenging. In Excerpt 10, two tutors ($2 and $4) and a student ($1) are talking about *bad conscience* which has been retrieved from the audio data they are collaboratively analysing.

Excerpt 10. ***bad conscience***

```
<$2>  Let's just step out of the data for a second and ask
      ourselves that question(.) If you have a bad conscience(.)
      how do you ↓feel(.)[Do you feel-
<$4>                     [It's almost like a rhetorical question
<$2>  =do= do you fee:l bad or good ↓though
<$4>  It's like a silly question
      Well I don't know= I'm asking gen= a genuine question(.)
<$2>  do you feel bad or good if you had a bad con↑science
<$4>  Well to me that sounds like a rhetorical question
<$2>  Can you answer it for us=
      [cos I am asking as a genuine question
<$4>  [Sounds like a silly question(.) How DO I feel if I have a
      bad conscience ((laughter))
<$1>  I wouldn't say yes uh- bad or good
<$2>  Well I mean if [you've done something BAD-
<$4>                 [I feel- probably feel bad if I've got a
      bad conscience
<$2>  =if you've= if you've= done something BAD(.) and you don't
      feel bad about it(.) does that mean you've got a good
      conscience or a bad conscience
<$4>  If I feel bad I'd probably feel pretty bad
<$2>  No no that's not what I'm asking(.) If you've= if you've=
      done something that you know is wrong-
<$1>  Ah I understand
<$2>  =↑Okay(.) You've= you've= you've hurt somebody ↑right
<$1>  Mm-hm
<$2>  And(1.0) should you feel good or bad ↑about it(.) and then
      if you feel good do you have a good conscience if you feel
      bad do you have a bad- it's not as straight↑forward as
      that(.) You would say somebody had a bad conscience in
      that example(.) if they(.) felt bad
```

In this excerpt, the goal of the participants is to make sense of a question retrieved from their data: "If you have a bad conscience, how do you feel?". The talk to examine the rationality of this question is initially co-framed by the two tutors when both of them choose to focus on the [FUNCTION] of the question. $4 claims that the question is like a rhetorical one since it seems to be asked by the speaker to produce an effect (e.g. to draw attention/elicit interest, to provoke thinking, etc.) or to make a point (e.g. someone should feel good/bad if they have a bad conscience, etc.). The pragmatics of the question, in such sense, is to motivate or

persuade rather than to pursue an answer. The question, however, is not well formulated as perceived by $4 to meet the purpose since it sounds like a *silly* one (Line 6) to the hearer with no further contextual information provided. By contrast, $2 tends to take the question as a *genuine* (Line 7) one (aimed to get an answer) and invites $2 to co-examine its answerability. Having failed to give an articulate answer, $2 reiterates that the question is silly (Line 12), whereas $1 enters into the talk by implying that the question might not be answered in a straightforward way. The co-framing initiation of $4, interestingly, receives a preferable response from the student instead of the other tutor. Nevertheless, it may not be fair to say that $2 does not respond to $4' co-framing initiation. His lexical choices to evaluate the question show subtle evidence of a conceptual shift. When taking the question as *rhetorical* he suggests that the answerability of the question is irrelevant in his script since its function is not to elicit an answer; while he immediately portraits the question as something *silly*, which implies that the question is almost unanswerable in the context. $2's responses show that his original framing tends to remain though the participants negotiate on the spot to achieve conversational cooperation (Gumperz 1982). The "resilience of schemas (frames)" (Tannen and Wallat 1986: 306) as such triggers $4's repetitions of his question and further elaboration on the object of the talk.

In the following talk, $4 tries to make himself clearer (*I mean*, Line 15) by suggesting that whether the question is answerable may depend on how one understands the meaning of *bad*. A conceptual process to make sense of *bad* can be identified from $4's successive lexical choices: *if you've done something bad* (Line 15) -- *if you've done something that you know is wrong* (Line 23) -- *if you've hurt somebody* (Line 25). One semantic potential associated with *bad* is made discursively relevant with $4's invoking the [MORALITY] frame to interpret its meaning. The lexical choices contribute to creating the discursive relevance by intensifying

the degree of "being bad": *bad* as a gradable adjective to be represented along a spectrum, *wrong* as a non-gradable adjective at the negative polar in the right-wrong dimension, *hurt* as a verb with a very strong negative connotation. The conceptualization becomes accessible to $1 (*Ah I understand*, Line 24), while $2 does not show whether he gets the point too.

The observation suggests 1) that the co-framing between the two tutors is insufficient if not unsuccessful throughout the talk even though they have a shared starting point to examine the data. While $4 keeps eliciting co-framing moves from $2, the latter fails to meet the expectation; 2) that when another participant who is not the expected co-framer makes co-framing moves, the original framer would probably create a new co-framing relationship with him/her by giving positive responses (e.g. ↑*Okay*, Line 25) and strengthening mutual understanding (e.g. *Mm-hm*, Line 26; *it's not straightforward as that*, Line 29-30); and 3) that lexical choices contribute to selecting the semantic potential of a particular lexical concept thus facilitate situated meaning construction.

5.4.3 *Other cases of framing*

Section 5.4.1 and 5.4.2 have already shown how alternate framings of a single situation and co-framings within/beyond the role boundary are applied by the participants to negotiate meaning in a situated way for knowledge construction. This section examined some examples of other cases in which the participants did frame their talk but used different devices from alternate framings and co-framings. The example I shall start with is Excerpt 11 in which a tutor is giving a short monologic illustration to a group of students of how discourse markers in a teacher's language use contribute to creating facilitative shared learning space. Prior to the excerpt, the tutor and the students discussed the characteristics of spoken academic discourse in university settings.

Excerpt 11. ***shared learning space***

<$1> …We =we found a lo:t of <discourse markers> of <shared space> (1.5) >Now let me just say a quick word about this< (.) Erm(3) Shared space means (1) er (0.8) how: do: (.)interactants create (0.6)the kind of space where we can do something usefully together (0.8)↑right (0.8) And one of the ways in which we do that (.) is we use words like "you know" (.) and "you see" (.) and "I mean" (.)and "kind of" (.) they're sort of va:gue language (0.5) and some people think (.) they are a sign of (.) disfluency (.) So in testing circles (.) if there was ↑too much of this (.)people would say this is disfluent speech ↑right (0.6) But in a teaching context (.) they're very important (.) ↑Without them (0.5) it sounds like a lecture (1.2) If you didn't hear me saying "you know" (0.5) and "as we saw last week" (0.5) and "kind of" er (0.5) I would sou:nd to you like I was lecturing you rather than (.) including you (.) So from a ↑learning point of view (0.8) creating shared space (.) is rea:lly rea:lly important (1.2) And if you: look at a piece =look at a transcript (.) and delete a:ll the discourse markers of shared space (.) >you'll see what I mean< it looks like somebody's actually (0.5) telling you rather than involving you ↑yeah (.) So they're really important (.)They're very (.) central…

The "learning" dimension associated with the lexical concept ACADEMIC (see discussion in Section 5.3.2) has been highlighted in the tutor's prior talk to define the boundary within which the core concept -*shared space* of learning- is relevant and prominent in the current discussion. By doing this, the tutor manages to organize the knowledge about systematic cognitive associations between concepts (e.g. *academic discourse, learning, shared space*) and uses this knowledge to direct interpretation for the learners (Tannen 1993). He goes on to frame the talk by introducing the concept of *shared space* as a visualized representation of where learning happens. The framing move is realized through his use of a metaphoric device. Metaphors are effective for meaning conveying since they express "an identity in structure between different domains" (Gärdenfors 2014: 39) thus perform as "mappings" to "preserve

the cognitive topology" between "the source domain" and "the target domain" (Lakoff 1993: 215). The lexical concept SPACE is conceptually transferred from the MEASUREMENT domain to the LEARNING domain, which creates new knowledge (Gärdenfors 2014) that learners are encouraged to take initiatives to create, manage and therefore shape "learning" through interaction. A perceptual shift of learner identity is expected as a result of this kind of knowledge construction, which could be potentially achieved by drawing upon certain activity patterns in the MEASUREMENT domain to understand the mechanism in the LEARNING domain.

Another framing move applied by the tutor is his explication on the mechanism of *shared space* of learning by drawing the learners' attention to a specific linguistic device: *discourse markers*. This makes it possible and feasible to evaluate "learning" in terms of features of talk. The lexical form *discourse marker* is introduced as the "informational focus" (Cowles et al. 2007: 4) to draw the learners' attention to the information which is not presupposed by the tutor's utterances but rather predicted of the main topic of his talk. This kind of cognitive representation concerns *dynamic* properties (Port and van Gelder 1995) of the LEARNING domain, typically represented by actions rather than static concepts like categories and properties (Gärdenfors 2014: 145). The tutor uses two groups of lexical forms - *lecturing* vs *including* and *telling* vs *involving* - to highlight the potential effects of discourse markers in negotiating the somehow ambivalent teaching-and-learning relationship. This further draws the learners' attention to the "contrastive focus" (Cowles et al. 2007: 4) concerning how to perceive the function of discourse markers. The emphases on *including* and *involving* are marked both structurally (with the conjunctional component *rather than*, Line 16) and prosodically (emphasized with a higher pitch and a higher volume, Line 16; 21-22).

Cognitive mappings which are metaphorically expressive as framing devices are not unique to Excerpt 11. Following are two other examples. Excerpt 12 (both $1 and $2 are tutors) shows how the source domain of TRAVEL is cognitively mapped onto the target domain of BECOMING A TEACHER, with the associations between the semantic elements interpreted based on the structurally transferrable conceptions (Table 26).

Excerpt 12. ***becoming a teacher***

1	<$2>	I'll come and sign a different PGCE of course= but
2		it'll↑ never be the ↓same (1) But the person ↑you're talking
3		about who we cannot name= names beginning with ↓D (1.2)o=
4		openly would say(.) "I've been at this school for <twenty-
5		five years>(1.2)Goodness knows >what I'd be like if I tried
6		to< start at ↑another school"
7		((24 lines omitted))
8	<$1>	>It ↑wasn't that we didn't have an idea of ↑who was good
9		at ↓what<(.)it's just that it wasn't <qui:te so explicitly
10		lai:d ou:t>(.)And ↑as soon as you lay: it ou:t(.)>you look
11		at it and you will think< "Ooh we should all be
12		↓outstanding"(.) because ↑why not(.) It's written down
13		there(.) and some people can be(.) So it's added the pressure
14		in(.)um(.) which I don't- nece= necessarily think is always
15		a ↑good thing(.) cos it makes= makes= people °quite
16		anxious°(.) ↑And also suggests that you are a finished
17		↓product(.)at the ↓end(.)you know that actually because
18		you've done a PGCE and you've got lots of nice "goo:ds"
19		and outstanding kind of comments(.)that= (that's it) you've
20		↓cracked it(.)but- you'll= you'll= never comple:tely crack
21		↑it

Semantic element	Source domain: TRAVEL	Target domain: BECOMING A TEACHER
Actor	Traveller	Student-teacher
Activity	Travailing	Being trained and practicing
Journey	Trip	Development
Source location	Starting point	As a student-teacher
Goal	Destination	To become a teacher
Path/area	Path/area	Instruments and resources
Co-participants	Fellow traveller(s)	Institutional counterpart(s)
Baggage	Luggage	Burden and responsibility
Support	Maps, route signs, compasses, etc.	Training, guidance, examples, feedback, evaluation, etc.
Impedance	Natural or physical obstructions	Physical, mental or social-institutional unfavourable conditions
Mode of transportation	On foot, by air, on the train, etc.	Methods and approaches

Table 26 Mappings between TRAVEL and BECOMING A TEACHER domains (ALE)

Note that what is emphasized in both of the source domain description and the talk in Excerpt 12 is the concept of PROCESS rather than BEINGINNING/END of the journey, which is explicitly expressed by $1's being opposed to taking student-teachers as finished products. The mappings are set up based upon cognitive projections from the source domain to the target domain in the participants' talk. The accessibility of the paths depends on the partially shared structures of the two domains (see Lakoff and Johnson 1980a, 1980b; Turner 1993). On the other hand, the metaphorical projections are found to be "locally creative" as "part of the ongoing reasoning and discourse construction" (Fauconnier 1997: 9). The localized creativity is revealed by the participants' deliberately bringing some of the frame elements into focus to construct knowledge by building up or re-conceptualizing their relationships through re-framings. Specifically, $2 focuses on the relationship between the years one stays in the teaching profession (the duration of the journey) and the degree of his/her passion and sentimentality (the traveler's curiosity and adventurous ambition). $1, while embraces the

emotional loading of beginners, focuses on the relationship between the application of standardized evaluation criteria (planned routes and recommended tourist destinations) and student-teachers' professional development (motivation to explore the unknown). The two tutors have different framing orientations and tend to construct the knowledge by attending to different frame elements. This indicates that interlocutors differentiate "information accessibility" (Sweetser and Fauconnier 1996: 2) by attending to the pragmatic selectivity rather than the semantic exclusiveness of lexical concepts.

Different kinds of mappings are involved in the participants' lexical choices, which can be traced from the collocational behaviors of words. For example, Table 27 shows the cross-domain mappings of the *web*-compound nouns through which the source domain of ENTITY is mapped onto the target domain of INFORMATION. The structure of the source domain is therefore used to think and talk about the structure of the target domain based on the mental projection (Fauconnier 1997a).

Second constituent stem in compound	Source domain (ENTITY)	Target domain	Mapping path
page	PAPER DOCUMENT	INFORMATION	Projection: Format
space	MEASUREMENT OF AREA	INFORMATION	Projection: Container
site	PHYSICAL LOCATION	INFORMATION	Projection: Location
flow	SAILING AND SWIMMING	INFORMATION	Function: Navigation
service(s)/server	PEOPLE OR OBJECTS	INFORMATION	Function: Demand
browser	PEOPLE OR ANIMALS	INFORMATION	Function: Purpose

Table 27 Cross-domain mappings of *web*-compound nouns (BI)

The cross-domain mappings shown in Table 27 are realized via different paths and are central to "understanding of semantic and pragmatic language interpretation and cognitive

construction" (Fauconnier 1997a: 12). Projection mappings (e.g. *web page*, *web space* and *website*) concern partially structural constructions of meaning while function mappings (e.g. *web flow*, *web service/server*, *web browser*) involve pragmatics that evokes the correspondence between two categories (Fauconnier 1997a) to produce textual coherence (Fillmore 1982).

Cognitive mappings are identified in NUCASE data not only across domains (e.g. metaphors) but also in a more complex configuration which contributes to partitioning discourse in different conceptual realms. This is the building and connection of mental spaces. As discussed in Section 2.3.1, mental spaces as conceptual structures are set up dynamically throughout the ongoing discourse based on linguistic and non-linguistic information (Fauconnier 2012) for interpretative purposes (Hamawand 2016) to realize local understanding and action (Fauconnier and Turner 2006: 307). Mental spaces are "internally structured by frames and externally linked by connectors" and make the discourse unfold through successive "cognitive configurations" (Fauconnier 1997: 38). Identifying mental spaces can help describe the semantic potential of lexical concepts encoded by particular linguistic forms, thus contributes to explaining how the real-world knowledge schematized within certain frames is organized and activated through one's lexical choices. In the following analysis, I shall focus on the cases in which lexical forms function as legitimate cognitive devices for mental space building and connection. Excerpt 13 is a talk scenario happening immediately prior to Excerpt 8. The biological expert (\$7) is explaining to the student research team the difference between a vaccine target and a drug target.

Excerpt 13. ***vaccine target and drug target***

<$7> So the ↑second one is sli:ghtly different from the first one(.) So with the first one you're looking for: a vaccine target(.) You're: essentially looking for molecules that are(.) specific to a particular(.)organism(.) that we can target as a vaccine
<$M> [Yeah]
<$M> [Yeah]
<$7> With the ↑second one(.) you're looking for: proteins that are similar(.) to proteins that are already to be known to be drug targets
<$M> Mm-mm
<$7> So: bear in mind that the two bits of the project(.) We're= we're= we're really looking for((coughing))different things
<$M> Yeah

To demonstrate clearly how the mental spaces are set up and connected, I present $7's talk into four sentences:

Sentence 1: *So the second one is slightly different from the first one.*

Sentence 2: *So with the first one you're looking for a vaccine target.*

Sentence 3: *You're essentially looking for molecules that are specific to a particular organism.*

Sentence 4: *With the second one you're looking for proteins that are similar to proteins that are already to be known to be drug targets.*

In this group of sentences, the base space is built up by $7 as associated with the [COMPARISON] frame which includes two elements *a* (the first part of the project) and *b* (the second part of the project). The lexical form *looking for* then sets up a new space of TARGET relative to the base space that partitions off the information about *a*. Sentence 3 sees the set-up of a new space of MECHANISM by the space builder *specific to*. Sentence 4 returns to build

upon the TARGET space by filling in the counterpart partitioned information about *b* and enters into the MECHANISM space. Note that the TARGET space set up from Sentence 2 is internally structured by the lexical concepts involved as a hidden space builder. Therefore another space of REFERENCE is set up with the referent projection of *a* as *diseases* while the referent projection of *b* as known *drug targets*. Figure 23 shows the mental space lattice.

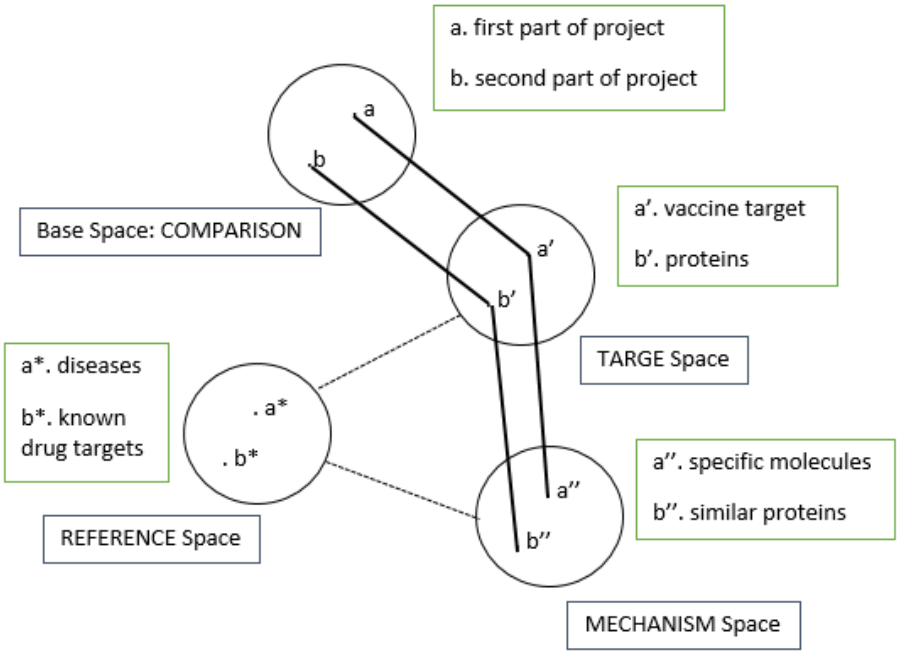


Figure 21 Mental space lattice in Excerpt 13 (BI)

Another example is shown below based on the talk in Excerpt 4 (Line 4-12), but with a more complicated structure (Figure 24).

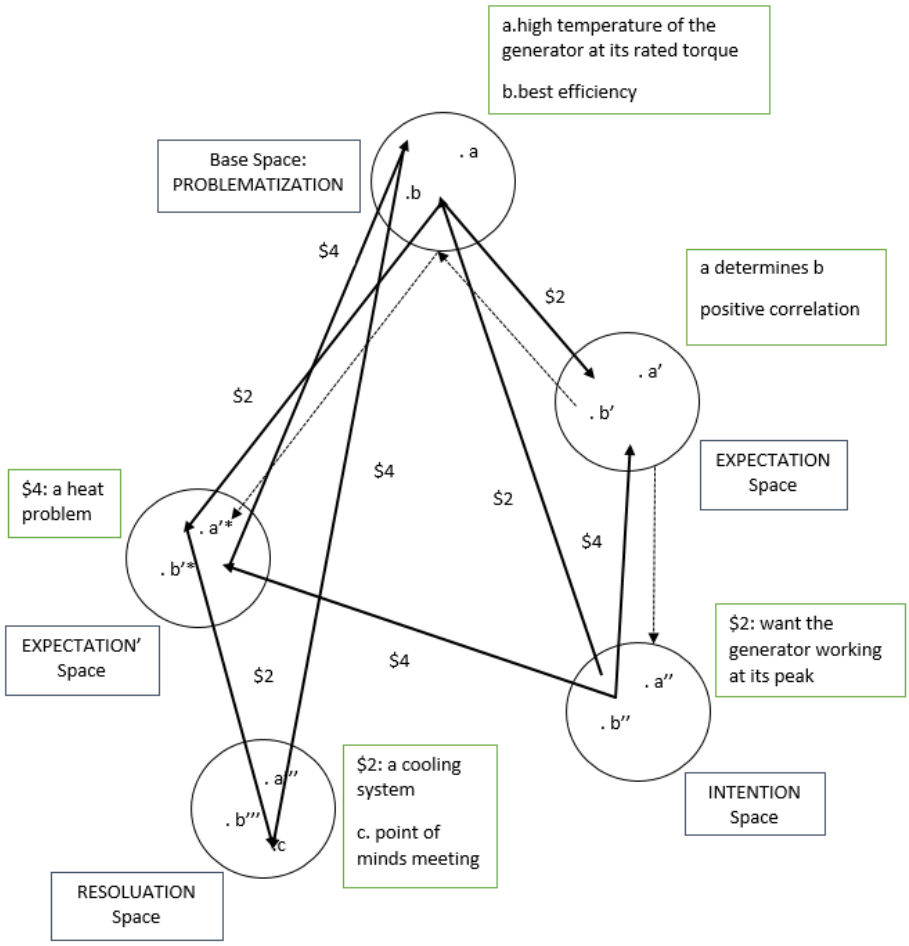


Figure 22 Mental space lattice in Excerpt 4 (ME)

It should be noted that the sequential analysis on the overlapping talk in Excerpt 4 (Line 10-11) has provided important discursive clues about how the mental spaces are set up and connected. The lexical form *but* right prior to the overlapping talk expresses the concept of CONTRAST which is mutually acknowledged. It is from here that the conceptual network sees another space projected from the EXPECTATION space which I shall call "a shadow space" of EXPECTATION'. The projection from the EXPECTATION space to the EXPECTATION' space experiences a detour which involves a shift of focus in the mapping path. The two participants' mappings are found to be divergently filling into the EXPECTATION' space, due to their different predictions of what *but* would evoke in the

following talk. Specifically, $4 draws on the contrast between preferred and dis-preferred effects of the generator's getting too hot, which leads him to return to the Base space once again. $2, on the other hand, focuses on the contrast between the problem and the solution, thus sets a new space of RESOLUTION. Note that these framing moves reveal the important role played by tense as space connecting devices in determining which space is in focus (Fauconnier 1997b: 41): the present perfect tense shows an anaphoric orientation while the simple present tense indicates a progressive concern. Without the detailed transcription of the overlapping talk, the researcher might risk taking $2's turn as a response to $4's turn, which would lead to a dramatically different representation of mental space building and connection.

The explication and visualization of the mental space lattices shown in the above two examples add to the evidence of human psychological 'priming effects' (Sweetser and Fauconnier 1996) which reveals a basic function of linguistic structures in communication to the "differential information accessibility that attends on cognitive viewpoint" (p. 2). We have seen the variability of the cognitive mappings for meaning negotiation in communication, and how the accessibility and selectivity of such mappings contribute to shaping the structurality of knowledge representation, transmission, and interpretation. The findings correspond to Fauconnier's (1992) claim that structural configurations of discourse are central to meaning production due to a dynamic relationship between discourse and knowledge (see also Cazden 1986; Cobb and Yackel 1996; Koschmann et al. 2000). The findings also indicate that framings through language use are observable and configurational and that invoking situated interpretation of framing is a powerful analytic tool to account for characteristics of talk-in-interaction.

Finally, we have cases in which shifting of framing behaviours happens due to the conflict between different ways of manipulating the motivating circumstances of knowledge construction. This kind of framing is related to what has been coined as *re-framing* by Fillmore (1982a) to illuminate various kinds of semantic change. Re-framing, in its original term, refers to "reconstituting the motivating circumstances while preserving the lexical item and its basic fit with the associated scene" (Fillmore 1982a: 387). However, the framing type I shall examine here has little to do with semantic change over time while the analytic focus is on how the participants' prior epistemic orientations, or in Goffman's (1959) term 'territorial reserves' of knowledge (see also Kamio's [1997] *territories of information*), bring personal facts and expectations into the moment-to-moment negotiation of meaning. Frame changes as such can be associated with the shifting relationship between the speaker and what he/she says (Lerner and Kitzinger 2007), thus are crucial to understanding communicative events (Kramsch 2002: 15).

As discussed in Section 2.2.3, knowledge domains which are disciplinarily defined often see a flatter gradient of epistemic status (Heritage 2012) among participants. This means that participants of the same discipline usually share the basic knowledge structures and skills as 'disciplinary facts'. Therefore, the potential of manipulating the motivating circumstances to re-schematize mutual academic experiences has been to a large extent limited. The observation is consistent with Heritage's (2013) claim that the epistemic status of participants for some domains of knowledge is "an easily accessed, unquestionably presupposed, established, real and enduring state of affairs" (p. 558). This may explain why this kind of framing is relatively rare in NUCASE when the participants talked about discipline-specific lexical concepts. Excerpt 14 is an example where two individual schematizations correspond to interpret a statement of a third participant.

Excerpt 14. ***a culture of classroom***

```
1   <$2>  I= I think the other thing= just one ti:ny thing on that
2         is= is the notion of sort of- how: a- a CUL:ture of a
3         classroom develops(.) So how: over time(.) depending on
4         what the tea:cher does in the classroom(.) the students
5         they will then lear:n to behave in certain ↓ways(.) So
6         for instance(.) if you do this(.) talk to the person next
7         to you before(.) having to offer a res↓ponse(.) you'll
8         ↓then when you don't do that you'll see people do it(.)
9         You know= so there's all sorts of you know= you've ↑SE:T
10        ↓up a set of-
11  <$4>  =Right(.)[Well kind of socialize them almost
12  <$3>           [↑YEAH(.) SOCIALIZING its ↑nor:ms its community
13        of practice-
14  <$1>  Yes
```

Prior to the beginning of Excerpt 14, the three tutors ($2, $3 and $4) and the students collaboratively analysed a transcript of spoken data. Their talk then progresses to seek the socio-cultural implications of classroom practice when $2 introduces the concept of *how a culture of a classroom develops* (Line 2-3) and points out how the development can be achieved using an example: The development is a long-term process (*over time*, Line 3) depending on the interaction between teaching and learning (*what the teacher does, the students will learn to behave in certain ways*, Line 4-5), in which the teacher plays a key role to foster the habitual norms of behaviour (*set up a set of…*, Line 9-10). $4 latches $2's talk (Line 11) by first providing an agreement and his following talk overlaps with $3's talk. $4 uses *socialize* (Line 11) to conceptualize what $2 has illustrated as the process of classroom cultural development. $3, taking a further step, explicates the social attributes of what $2 has incorporated into the process: behaviours as *norms* (Line 12), a classroom as a *community* (Line 12) and teaching and learning as *practice* (Line 13).

Note that the overlapping talk in Line 11-12 is more appropriately seen as two personal activations of the conceptual knowledge agreeing with each other than a co-framing phenomenon. This is because:

1) The overlaps begin early right after $4's agreement token *right* (Line 11). The lexical unit *well* (Line 11) at which the overlaps happen does not convey any initiation of co-framing from $4;
2) The lexical unit at the initial position of $3's turn *yeah* (Line 12) is uttered with a rising intonation, which may indicate that it is a delayed response to $2's talk rather than a follow-up of $4's *right* since it might be used as a preliminary to something that follows in a multi-unit turn; and
3) $4 and $3 use exactly the same lexical form *socialize* almost at the same time to conceptualize $2's talk, which indicates that the meet-of-minds is achieved before they utter the word rather than negotiated in interaction. In addition, they orient to completing their respective turns even after hearing the other's utterance of the word *socialize*, which may suggest individual framings rather than co-framings.

Individual schematizations, however, do not always correspond and can be put in contrast. Excerpt 15 includes a talk scenario happens immediately prior to Excerpt 10. Three tutors ($2, $3 and $4) and two students ($1 and $6) are talking about a student's answer ("*Bad*." with laughter) to the teacher's question ("*If you have a bad conscience, how do you feel?*") in the audio recording.

Excerpt 15. ***the student's answer***

1	<$1>	Cos like- you are right in a point that- in line- th=
2		thirty-three(.) he asks(.) um(.) "Okay if you have a bad
3		conscience how do you feel(.) ↑Bad"(.) I mean-
4	<$3>	((laughter)) =That's a ↑completely NORmal answer(.) How=
5		how do you feel ↑bad
6	<$4>	°It's almost undermining the ↑tea↓cher's(3.0)↓job that
7		you're almost using it(.)for humorous effect°
8		((3 seconds of silence))
9	<$6>	Yeah
10	<$2>	Well yeah [I mean-
11	<$4>	[You're exploiting(.)the teacher's-
12	<$2>	↓But(.) simultaneous↑ly(.)it's a PER:fectly rea:sonable
13		answer(.)>If you have a bad conscience how do you
14		feel?< °You feel ↓bad°(.) It's an ENTIRELY reasonable
15		answer-
16	<$3>	But maybe even-
17	<$2>	=But because of the con↑text(.) because of what the
18		teacher's trying to do pedagogica↑lly(.) he's like- ↑I
19		think >there's a couple of things going on here<(.)
20		[He's= he's-
21	<$1>	[There's also laughter
22	<$2>	Yes(.) There's the laughter which is= is-
23	<$3>	=It could be an orientation how- s= strange a question it
24		is-
25	<$2>	↑Yeah
26	<$3>	=The wha= what answer can you give there=
27		[that's a satisfactory answer-
28	<$2>	[Absolutely(.) I= I think your po= your point's really
29		strong there

Prior to Excerpt 15, $1 as the one who transcribed the data has confirmed $6's argument that the teacher's question seems to be problematic with no contextual information provided to his students. That $3 latches $1's talk with laughter (Line 4) again indicates the absurdity of the question. Here $3's stance towards the student's answer is brought into the talk based on his evaluation of the latter from the question's point of view. In other words, the motivating schematization of $3's interpretation is the communicative relationship between a genuine

question and a legitimate answer. From such a perspective, the student's answer can be seen as completely normal because 1) it answers the question and 2) it performs a preferred next action (*bad* as an adjective to describe "how" one feels).

A shift of framing happens in the following talk with $4 entering into the interaction by highlighting the consequence of the student's answer: *almost undermining the teacher's job* (Line 6); *for humorous effect* (Line 7). A disapproval attitude is adopted by $4 to imply that the student may not take the teacher's question seriously and tend to use the answer to obstruct teaching or attract the peers' attention. The motivating schematization of $4's interpretation is the power relation between the teacher and the student in the classroom, in which the pedagogical goal is ultimate and the student is obliged to behave in a co-operative manner. On the other hand, $4 may orient to downgrading his epistemic veracity to display that he is less than fully committed to his talk (Park 2012). This is marked by his use of *almost* twice as hedges, a long intra-turn pause (3 seconds), and remarkably lower volume of talk. What is also noticeable is the prolonged transition space (3 seconds of silence) after $4's turn (Line 8). As an action (a statement) has been completed at the end of $4's turn, the silence is not attributable to any particular participant not speaking at the moment (Liddicoat 2007) but may indicate a lapse in the talk for the fellow participants to process the shift of framing initiated by $4.

$4's move to secure his framing is conducted after $2 and $6's minimal acknowledgements and overlaps with $2's following talk. $2's use of *I mean* (Line 10) may project a contrasting move with his acknowledgement (*well yeah*, Line 10) since the lexical form can be used to correct what one has just said. The contrasting orientation is confirmed by the latched *but* (Line 12) when $2 regains the floor after the overlapping talk. The *but* at the initial position of

$2's turn not only demonstrates a cognitive level contrast (Blakemore 2000; 2002) between the distinct framings of $3 and $4 but also has conditional relevance (Schegloff 1972; 1992) to $4's maintaining of his framing as the first action. The *but* at the sequential level initiates a dis-preferred next action from $4's perspective, while shifts the focus back to $3's framing which is constructed in the prior course of action (Hata 2016).

$2's talk (Line 12-15) enhances $3's framing by reiterating that the student's answer is *perfectly* and *entirely* reasonable until $3 initiates a possible contrasting action with his original framing (the turn-initial *but*, Line 16). This triggers another "first action – next action" sequence in which $2 latches $3's talk when the contrasting orientation of the latter becomes mutually accessible (the latched *but* at the initial position of $2's turn, Line 17). $2's following talk (Line 17-20) then shifts the focus to $4's framing by highlighting the pedagogical considerations for the interpretation of the question-answer pair. The re-focus on $4's framing, however, does not maintain when $1 mentions the laughter in the student's answer (Line 21). This triggers $2 and $3 to focus on $3's original framing again in which the teacher's question is portrayed as *strange* (Line 23) and the student's answer as *satisfactory* (Line 27).

The observation shows that distinct framings can be applied based on different ways of schematizing one's prior knowledge and the degree to which such schematization is sensitive to the local context. For example, in Excerpt 15, $3's schematization is more sensitive to the on-going interaction than $4's schematization: the power relation between teachers and students tends to be more stabilized and de-contextualized than the pragmatic relationship between a question and an answer. In other words, $4 is more schema-oriented while $3 is more communication-oriented. The contrastive framings compete to construct situated

meaning throughout the talk with the participants' focus shifted between them. At the end of the session (talk after Excerpt 15), the participants achieve the meet-of-minds by successfully identifying the connection between the two distinct framings: the teacher's unsuccessful pedagogy makes the question conceptually difficult to answer, which causes an invalid student contribution to learning. In this way, the knowledge is constructed around the issue of how teaching practice in the classroom can be effectively associated with certain pedagogical goals.

In sum, Section 5.4 examined framing sequences for knowledge construction in the participants' small group interaction. Expanded talk scenarios were examined around the target lexical concepts which had been initially analyzed in the prior two sections. Interactional Linguistics techniques were employed to conduct close, detailed observations of the participants' framing moves at the sequential level. The analyses suggest that the participants applied various framing devices to construct knowledge through situated meaning negotiation. Alternate framings were found to be applied by the participants when a single situation was the focus of the talk. Different ways of framing a single situation can be a result of contrasting actions (Excerpt 1), different viewpoints (Excerpt 2), distinct institutional values (Excerpt 3), conceptual mismatches (Excerpt 4) and management of framing mutability (Excerpt 5).

Compared with alternate framings, co-framings showed a closer association with speaker roles which were either previously assigned or naturally emergent from the talk-in-interaction. Participants usually tended to co-frame the talk when they were with a shared orientation to a specific task (Excerpt 6) or with similar roles which were institutionally specified (Excerpt 7). On the other hand, co-framings beyond the role boundary were also identified to be applied,

with individual framing moves showing conditional relevance to the prior courses of action and negotiation of epistemic stance showing reverence for more powerful social groups (Excerpt 8). More complex structures were expected to be associated with co-framing behaviors when the listener's interpretation needed to be adjusted to the change of element(s) in the speaker's framing (Excerpt 9). However, co-framings can be challenged thus risk failure in some situations where a selected co-framer does not live up to the expectation of the co-framing initiator or refuses to adjust his/her interpretation when the former changes his/her representation of certain frame elements. Specifically, the sequential projection of possible contrasting actions (Excerpt 8) or simply the conceptual complexity of the topic (Excerpt 10) can cause co-framings to fail in interaction.

Some other framing devices were also found to be used by the participants, which included cognitive mappings across domains (Excerpt 11 and 12), mental space building and connection (Excerpt 4 and 13), and re-schematization of motivating circumstances (Excerpt 14 and 15).

5.5 Summary

This chapter examined how the NUCASE participants framed their small group interactions to collaboratively construct knowledge through the talk of discipline-specific lexical concepts. Specifically, Section 5.2 gave an overview of the key semantic domains in each sub corpus and presented examples to illustrate how lexical concepts encoded by linguistic forms provide access sites to coherent conceptual knowledge structures. This section provided evidence of typical semantic frames evoked by linguistic forms across NUCASE. The findings suggest that meanings associated with linguistic forms are not calculated independent from context; rather, they are mentally represented by particular lexical concepts encoded by the linguistic

forms thus perform a function of language in use.

Section 5.3 focused on how interpretive frames were invoked by the participants to assign the target discourse relevance and coherence for situated meaning-making. Individual framings were investigated at the discursive level to show how particular knowledge structures were accessed and activated to make the talk mutually understandable. The findings suggest that the interpretive frames invoked by the participants are partial representations of larger-scale conceptual knowledge structures. The semantic potential of lexical concepts provides access to the structures and determines what aspects of the structures are to be activated for local meaning negotiation. It is argued that the act of invoking interpretive frames is highly context-dependent and goal-oriented thus is of great importance to maintain the basic mode of institutional interaction.

Section 5.4 examined framing sequences within expanded interaction of the participants' small group talk. The findings suggest that the participants applied various framing devices to construct knowledge, including alternate framings of a single situation, co-framings, cross-domain mappings, mental space building and connection, and re-schematization of motivating circumstances. The duality of framing was identified: the stability of what occurs as meanings which are shaped by the participants' prior knowledge and the variability of interaction which makes meanings emerge from context rather than being given in advance (see Tannen and Wallat 1987). It is argued that pragmatics are fundamental to framing through lexical concepts even though the latter represent the semantic pole of linguistic forms (Evans 2006) as stored in speakers' knowledge of language use. It means that there is an interaction between cognitive configurations of frames and framing moves at the action level. This highlights the importance of conducting sequential analyses of framing-in-interaction to examine the ongoing process of meaning negotiation and knowledge construction.

Chapter 6. Discussion

6.1 Introduction

Combining Corpus Linguistics and Interactional Linguistics approaches to the NUCASE data, this study has explored how the participants framed their small group interactions to collaboratively construct knowledge through the talk of lexical concepts. More specifically, in the preceding chapter, the following has been analyzed: 1) how the lexical concepts encoded by particular linguistic forms provide access to coherent, non-linguistic knowledge structures; 2) how the participants invoke interpretive frames to assign relevance and coherence to the target discourse; 3) how the participants apply various framing devices in interaction to construct knowledge through situated meaning negotiation. The observations suggest that:

- Meanings associated with linguistic forms are not calculated independent from context; rather, they are mentally represented by particular lexical concepts encoded by the linguistic forms thus perform a function of language in use.
- The interpretive frames invoked by the participants for situated meaning-making are partial reflections of larger-scale conceptual knowledge structures. The semantic potential of lexical concepts provides access to the structures and determines what aspects of the structures are to be activated for local meaning interpretation.
- The participants applied various framing devices to construct knowledge through situated meaning negotiation in their talk. The duality of framing is revealed by the stability of what occurs as meanings shaped by the participants' prior knowledge and the variability of interaction which makes meanings emerge from context.

This chapter re-examined the observations presented in the preceding chapter and discussed the research findings in relation to the literature (Section 6.2). Shortcomings of the research design were acknowledged through methodological considerations (Section 6.4). Recommendations for further research were given at the end of the chapter (Section 6.5).

6.2 Discussion of research findings

6.2.1 *The institutionality of university small group talk*

The interaction scenarios in NUCASE are found to see a discursive shift of the interactional dynamics due to more active participation of learners. A resistance to "the teacher's agenda" (Smagorinsky and Fly 1993: 159) along with the learners' raised awareness to draw on their own identities, ideas and expectations in active meaning construction has contributed to the shift. In addition, university students as adult learners are expected to employ learning strategies and social resources at higher cognitive levels. Some may also argue that when the language is not taken as the learning object (as what has been done in language classrooms), knowledge construction would be better facilitated by meaning negotiation in a smoother way when the production of correct language forms becomes irrelevant or less emphasized.

While all the above propositions may be rational, they have to a large extent overlooked the underlying factor that yields the shift and defines its pattern; namely, the institutionality of university small group talk. This study takes university small group talk as a discursive variant of "institutional interaction" (Drew and Heritage 1992: 3), where participants' professional identities are made relevant to their working practices. Table 28 summarizes the basic characteristics of institutional talk with an example in NUCASE for illustration.

Characteristics of institutional talk	Example in NUCASE: Excerpt 8 and Excerpt 13
The interaction involves goal orientations which are tied to relevant institutional identities	Goal: targeting a vaccine strain; Identities: university student research team vs. expert in drug company delegation; Talk fashion: a top-down process with convergent contributions to the goal based on a shared understanding of task nature.
The interaction involves special constrains as an allowable contribution to the business at hand	Participants must enact the consultancy meeting as planned. No departures are allowable to be deviant from the agenda. The functionality of the institutionality lies in an integration of reference to the formal character of interaction and localized negotiation of understandings (Drew and Heritage 1992).
The interaction involves special inferential frameworks and procedures particular to specific institutional contexts	The expert as a more professional actor withholds expressions of elicitation, suggestion, and evaluation; with certain institutionalized claims to superior knowledge (Mehan 1985). This defines the scope of reasoning and expectations in the interaction.

Table 28 Characteristics of institutional talk represented in Excerpt 8 and Excerpt 13 (BI)

Drawing upon the inspirations from the discussion of institutions in the field of education (e.g. Erickson and Shultz 1982; Mehan 1993), university small group talk in NUCASE is found to show structural uniqueness. On the one hand, it does not resemble mundane conversations in that it is much more institutionalized showing "a heightened use of procedures" (Edwards and Westgate 1994: 116). On the other hand, it differentiates itself from traditional classroom interaction in that it features an equal participation among interlocutors and more emergent turn-taking patterns with pedagogical orientations less relevant or salient. In such sense, it becomes quite ambivalent to say that the prime goal of university small group talk is to maximize learning from the teacher/tutor's side, since the "goal-orientedness" is more often involved in interactional processes of meaning negotiating, extensive discussion and problem-solving, and knowledge construction is often mediated at higher cognitive levels.

The asymmetrical power distribution across different identities is the very reason why the knowledge structure of university small group can be maintained and managed. As suggested by Linell and Luckmann (1991), the underlying motive of human communication is the product of relevant inequalities of knowledge or the relative access to particular epistemic domains (Heritage 2012). Such inequalities of knowledge, however, do not necessarily mean that participants are at different levels of cognitive development or professional expertise, but refer to the degree of limitation in their access to available communicative resources (see Moje and Shepardson 1998) for knowledge production, construction, and transmission. It is the "restrictions on the partition rights" imposed by the "organizational routines and procedures" (Drew and Heritage 1992: 49) that shape the moment-to-moment representations of individual knowledge orientations.

In addition, divergences between one's epistemic status and stance (Raymond and Heritage 2006) may be strategically deployed to take advantage of interlocutors' resources. For example, the analysis of one small group talk scenario in NUCASE shows that members of the company delegation at a consultancy meeting persist in maintaining an epistemic stance that privileges the student's research team in terms of the access to the knowledge of a software design. By appearing less knowledgeable than they really are (Heritage 2012), the delegation aims to elicit as much information as possible from the research team so that the latter's professional qualification can be verified.

The above discussion contributes to a better understanding of learners' engagement in meaning negotiation for knowledge construction (Kittleson and Southerland 2004). First, there is a shift of motivation mechanism in university small group talk that it is the task, not

the pedagogy that makes the learners involved, interested and concerned. Teachers or tutors do not have to explicitly transmit procedural information relying on a certain language mode since the information is usually implied by the task nature and expected to be mutualized through effective communication. The group tasks in NUCASE involve minimum pedagogical orientations, while having direct associations with the real world at both the cognitive level and the practical level. Second, meaning negotiation is better facilitated and opportunities of knowledge construction taken when linguistic and communicative resources become more accessible in a more liberal learning environment. This is not to say that students in other learning settings are not capable of drawing upon resources to communicate, but that the range of their choices might be to a large extent limited by the institutionality of the setting. For example, in typical L2 classrooms, most moves in procedural context interaction are similar and so intensively loaded with pedagogical goals that the teacher and students can only "operate within a narrow range of language, much of which is formulaic" (Abdesselem 1993: 229; see also Seedhouse 2004b).

Previous studies have yielded empirical insights into the dynamics of working practices in institutions (Sarangi and Roberts 1999), among which the significance of constructing institutional knowledge has seen a growing awareness among scholars (see Douglas 1986; Mehan 1993). The previous chapter has already shown that framing moves can be closely associated with speaker roles and discursive nature of the institutional talk, while the underlying impetus is worth more attention. According to Olsen (2007), "the degree and form of institutionalization impact both *motivation* and *capacity* to follow institutionalized rules and codes of behavior" (p. 5, original emphases). The institutional knowledge, as perceived in this study, refers to both WHAT to talk and HOW to talk in university small groups. I argue that the content of talk and the ways of talking are *motivated* by contextual relevance and

realized through the participants' communicative *capacity* represented from their framing practice.

The institutionality of university small group talk can be interpreted procedurally through discursive framing. The application of cognitive frames and framings-in-interaction corresponds to the fundamental mechanism of institutional interaction: the way group members use 'categories' for sense-making. Categories, as the essential resources in institutional practice (Makitalo and Saljo 2001: 59; see also Douglas 1986; Bowker and Star 1999; Sarangi and Slembrouck 1996), relate to invoking structures (see Schegloff 1991, 1992) to understand talk-in-interaction. This is theoretically and methodologically compatible with the perspective of framing for knowledge construction. It suggests that conceptual constructions are realized internally rather than assigned externally, drawing upon the aspect of context being *procedurally consequential* (Schegloff 1992) in a discursive organization.

The claim that institutional characteristics at the linguistic level can be derived from the sequential organization of talk without being attending to "any information beyond the words on the page" (Heritage 1984: 283), however, has overlooked the role of motivating context (Fillmore 1982a) from an analytic perspective. As argued by Hester and Francis (2000), this perspective "disengages sequential structures from the situated identities or membership categories of the participants" (p. 497). The motivating impact of institutional contexts is thus expected to be instrumental to the interactional dynamics of university small group talk. The relationship can be articulated (refined from Olsen's 2007 proposition) as: the institutional routines and procedures play a crucial role in creating the motivation of participation and define the scope of accessibility of communicative resources which linguistically represent participants' capability in social interaction.

6.2.2 *Lexical concepts, meaning-making, and duality of framing*

This study has addressed the role of frame and framing in knowledge construction. Based on the observations in the preceding chapter, I argue that semantic and pragmatic constructions of meaning are inseparable in inquiries of knowledge construction through communicative language use. This usage-based perspective (Croft 2000; Langacker 2000) corresponds to the view that lexical meaning retrieved from human interpretation is highly flexible, open-ended and context-dependent (Fillmore 1982a, 1982b; see also Langacker 1987; Clark 1996; Sweetser 1999; Croft 2000; Fauconnier and Turner 2002; Tyler and Evans 2003). That "words are contextual expressions" (Evans 2009: 22) indicates that meaning is *assigned* rather than *carried* by associated words in utterances and such assignment is by no means unambiguous but "in service of the expression of situated communicative intentions" (Evans 2009: 22). In other words, meaning is "not a function of language *per se*, but arises from language use" as "a property of situated usage-events" (Evans 2006: 491).

As meaning negotiation is posited at the centre of knowledge construction (see Nunan 1988; Pica 1994; Scardamalia and Bereiter 2006), inquiries of the role of language in such processes require the search of theoretically legitimate and methodologically traceable units of meaning. The search for units of meaning has to do with linguistic forms conventionally associated with semantic units. While semantic units are constituted of lexical concepts linguistically encoded by words, meaning is not a property of words but represented within conceptual knowledge structures (Evans 2006). Such conceptual knowledge structures are called *frames* in this study. Frames as discursive constructs are introduced for an approach to language as a communicative system, providing a way of characterizing the connection between linguistic and non-linguistic knowledge (Cienki 2012). A lexical concept differs from a frame in that the

former is not a unit of meaning but has a meaning potential for situated interpretation while the latter has a coherent meaning of what knowledge is in it and how the knowledge is to be used thus can be taken as a coherent unit of meaning.

Lexical concepts are associated with frames because elements of a frame are "encoded and understood linguistically" (Telles-Ribeiro and Hoyle 2009: 77). Being conceptual in nature, lexical concepts are perceived to relate to "knowledge structures which are specialized for symbolic representation" (Evans 2006: 508). They are linguistically encoded mental representations of what we mean by "knowledge". As discussed in Section 2.3, frames are also used for knowledge representation to mediate the response of human thinking system to real world experience. A frame, however, is not directly related to language use by its origin but aimed to explain how patterns of mental activities structure common-sense thought (Minsky 1974). Nevertheless, as language and thought cannot be separated in the dynamic process of meaning construction to make sense of the world, the scope of a frame is expanded beyond the cognitive level to the communicative level. The relationship between language and knowledge representation is therefore made articulate that linguistically represented meaning can be characterized in terms of schematized human experience (Petruck 1996). A frame in such sense is taken as a system of linguistic choices associated with prototypical scenarios (Fillmore 1975). It is the linguistic choices that connect lexical concepts to frames since lexical concepts are inherently linguistically encoded and form-specific (Evans 2006) and frames provide conceptual structures within which the encoded linguistic information is processed, categorized and transmitted for communicative purposes.

The next step is to figure out how frames as units of meaning can be methodologically traceable in a principled collection of language data. The design of this study owes a great

deal to Sinclair's (1996) revolutionary view that it is not a single lexical word but its co-occurrence tendencies with other words and word classes that should be taken as valid evidence of how meaning is constructed in naturally-occurring discourse. Characterization of discourse based on the co-selection principle reveals human reliance on structures to communicate. This rationalizes the examination of collocational patterns of specific linguistic forms using Corpus Linguistics techniques. The frequency value of lexical units or sense categories as a retrievable and comparable indicator makes it a useful signpost to direct analysis from the immediate text into the surrounding context, in which way it can portray a model of discourse to highlight specific language features (McEnery 2015).

Collocational tendencies of lexical units are represented in the general patterns of selection determined by the unique lexical profiles of lexical concepts (Evans 2006). Lexical concepts define structural dimensions of words' textual statuses at "the semantic pole of a symbolic unit" (Evans 2009: 148). A lexical concept, as the most basic theoretical construct for human cognition and a basic unit for knowledge representation (Clausner and Croft 1999), suggests "semantic potentials" of words (Evans 2009: 69) to be conveyed and interpreted in situated ways. Semantic potentials facilitate access to distinct sets of conceptual structures (Lowe et al. 1997), namely, frames evoked by particular linguistic forms or invoked from language users' mental grammar (Hamawand 2016). Frames invoked for situated interpretation of discourse, however, are not conceptually accessed indiscriminately across contexts. The research findings presented in the preceding chapter (see Section 5.3) suggest that the interpretive frames invoked by the NUCASE participants for knowledge construction are partial representations of particular frames. The activation of conceptual knowledge structures does not usually happen on a large scale but quite incrementally in the on-going interaction.

The findings contribute to a better understanding of the distinction between the representation of knowledge in long-term memory and the representation of knowledge in on-line comprehension in communicative contexts. While some conceptual structures are fully accessible constituting a large amount of knowledge about particular topics, only a small portion of that knowledge may be or can be available at a particular moment of interpretation (Coulson 1997) for particular interlocutor(s). In such sense, situated meaning negotiation is an active process in which interlocutors integrate pre-existing knowledge external to the immediate encounter and on-going aspects of conceptual information to make sense of the discourse they are engaged in. Building partial representations of frames is, therefore, a defining feature of dynamic knowledge construction in that it not only retrieves frames to facilitate language compression (Coulson 2006) but also re-organizes frame elements for situated meaning-making. The processing difficulty of a given lexical concept associated with a particular frame thus depends on its predictability in the context where it is brought to be semantically and pragmatically relevant.

Understanding properties of lexical concepts also contributes to characterizing the target discourse if the researcher intends to reproduce the scenes of meaning-making from the insider's (the participants of the target discourse) view. The research findings in the preceding chapter (see Section 5.2) suggest that linguistic forms are not used in discourse by virtue of their semantic senses calculated independent from context; rather, they are chosen based on shared meaning represented by their associative lexical concepts. As characterized by Evans (2009), lexical concepts on the one hand "encode linguistic content" and on the other hand "facilitate access to conceptual structure" (p. 140). The bipartite structure is conceptualized beyond the "form-function" dichotomy, creating a relational dimension (Barsalou 1992)

where a semantic network is constructed to address the complex relationship between schematized knowledge patterns, selective tendencies, and online linguistic choices.

In the relational dimension, lexical concepts are related to each other one way or another to highlight specific aspects of particular frames for situated meaning-making. For example, the "teaching" aspect of the [EDUCATION] frame in ALE sub corpus is emphasized over the "learning" aspect with the "subject slots" of the frame filled with linguistic forms representing teaching roles and activities. The lexical concepts encoded by the word *repurpose* in BI sub corpus are mapped onto the structure of the [INTENTION FOR ACCOMPLISHEMENT] frame evoked by the word *purpose* for contextualized interpretation. Lexical concepts encoded by relational verbs in the [MOVEMENT/TRANSPORTATION BY WATER] frame in ME sub corpus represent different patterns of human intervention in the use of water, which provide access sites to different frame elements.

The observations in the previous chapter suggest that lexical concepts serve to modify conceptual structures (Evans 2009), through the projection of human cognitive propensity into a discursive practice which is thought-provoked, experience-grounded and inherently informational (Hamawand 2016). The linguistic information imported in communication, in such sense, neither contains nor conveys meaning; rather, it prompts local meaning construction by competent language users (Coulson 2006). This concerns a re-conceptualization of the traditional models of language comprehension by highlighting the creativity of knowledge construction mediated by framing in discourse. The speaker uses language to activate part of particular frames in the hearer's schematized knowledge base (Lee 2002) and then based on mutual expectations the speaker and the hearer construct an

abstract structure to conventionalize the sequences of communication by defining the roles, purposes, and routines of actions (Fillmore 1982a, 1982b).

Framing is a process which involves an interaction between human cognitive activities and communicative needs. The duality of framing defines iterative rather than linear logical sequences, through which language users frequently go back and forth evaluating contextual factors to interpret meaning represented in the knowledge structures (Coulson 2006) which motivate relevant linguistic categories into the speech situation. The application of cognitive frames as relatively stable knowledge structures always involves "contextualizing or situating events" in ongoing discourse to establish, negotiate and maintain "patterns of framing" (Fillmore 1982a: 391) at the action level.

The duality of framing is found to be closely associated with human experiences (Telles-Ribeiro and Hoyle 2009), expectations (Fillmore 1982a) and interpretation of meaning (Tannen and Wallat 1993) involved in social interactions. Detailed analyses of the small group talk sequences using Interactional Linguistics methods in this study contribute to uncovering the relationship between knowledge schemata and interactive framing behaviours. The research findings (see Section 5.4) suggest that the construal of prior knowledge as schematized conceptual structures do not necessarily constrain the creativity involved in meaning construction (Coulson 2006) but serve as contextual resources (Linell 1998a) which are sensitive to the on-going interaction. Such contextual resources are mediative, projecting what the participants have already assumed, believed, known or understood about what is talked about thus contribute to building a model of discourse (Linell 1998a). The model of discourse further defines how speaker roles are to be assigned or emergent following a particular procedure of actions. A network of interpersonal relationship is constructed this

way in the group, within which the participants are prompted and mutually facilitated to claim and negotiate their epistemic stances (Heritage 2012).

Curiously as shown in this study, however, divergent framing orientations do not often emerge because the participants hold distinct viewpoints of what particular knowledge is; rather, they are developed from different ideas of how particular knowledge is to be interpreted to complete specific tasks. Ways of framing are found to be closely associated with institutional goals and action routines. Specifically, alternate framings around a single situation usually happen in the general discussion of small group talk where different conceptualizations of a particular topic or event are tolerated or even encouraged for the purpose of information exchange and individual meaning representations. By contrast, co-framings usually happen when at least one group of the participants involved in the talk is highly goal-oriented, for example, to give instructions, to explicate working procedures, to produce extended explanations, to provide evaluative comments, etc. In addition, co-framings are also frequently conducted when the participants try to make sense of a complicated concept collaboratively.

Some framing devices are not employed to introduce new knowledge but to construct conceptual structures to facilitate mutual understanding, such as the use of metaphors based on cross-domain mappings and building and connecting mental spaces for partitioning of discourse. Some framing devices, on the other hand, involve the participants' endeavours to shift each other's framings (see Coulson 2006), such as the act of re-schematizing the motivating circumstances of the frames which have already been invoked in the interaction. Contrastive framings are conducted along with a more intense competition for speakership throughout the talk. Note that such kind of framing is quite rare in NUCASE when the

participants talk about lexical concepts which are highly discipline-specific. This may be because a disciplinarily defined epistemic domain usually involves relatively established epistemic status and stance among its participants (Heritage 2013), which means that the relative access to particular knowledge is more or less settled in the large bulk of interaction (Heritage 2012: 6). Therefore, the potential of manipulating the motivating circumstances to re-schematize shared academic experience has been to a large extent limited. The following section will specifically focus on the relationship between framing and the structure of knowledge construction discourse in NUCASE.

6.2.3 *Knowledge construction discourse*

The university small group talk examined in this study can be categorized as *knowledge construction discourse*, in the processes of which participants solve problems and construct knowledge of concepts, phenomena, and situations collaboratively relying on their language use (van Aalst 2009). Individual sense-makings through meaning negotiation are involved, leading to effortful, situated and reflective practices of learning (Palincsar 1998). The knowledge construction discourse used in the academy (institutions such as universities), according to Hyland (2011), "simultaneously constructs the social roles and relationships which create academics and students and which sustain the universities, the disciplines, and the creation of knowledge itself" (p. 171).

The knowledge construction discourse takes knowledge as a mental state of understanding something arrived at through constructive processes involved in people's making sense of it (Chang-Wells and Wells 1993). People come to know what is called "knowledge" by being engaged in a broad range of activities purposefully and meaningfully. In practice, most activities require a transaction between knowledge that is propositional and knowledge that is

procedural; or in Ryle's (1949) words, "knowing that" and "knowing how to". The former is tied to specific epistemic domains, while the latter consists of routines, procedures and strategies constructed through personal engagement and inter-personal communication across different epistemic domains (Chang-Wells and Wells 1993).

According to von Glasersfeld (1995), the basic assumption of the process of knowledge construction is that participants must make their ideas relevant and meaningful to their prior knowledge and the contextualized situation they are currently engaged in. The prior knowledge can be obtained from formal education or informal life experience. It is then to be shared in terms of the content (relevance), the amount (sufficiency) and quality (knowledgeability), the social environment where it is shared (power and competitiveness), and the specific sharing practices to be cultivated (means and boundaries) (Lencioni 2002). In such sense, knowledge sharing is an essential part of knowledge construction discourse. This rejects van Aalst's (2009) proposition that there is a lack of attention to interpretation, evaluation, and development of knowledge when it is shared as information. Knowledge sharing is indeed reflective because it reflects the sharer's orientation and projects the possible ways in which the receiver would interpret the sharer's intention thus to process, evaluate, develop and further share the knowledge.

In knowledge construction discourse, participants rely on structures of personal and collective knowing through socially situated learning biographies (Rommetveit 1985). The structure of sense-making practice determines the structure of discourse where knowledge is to be constructed. Based on the research findings in this study, I argue that for Higher Education that manifests features of the *integrated knowledge codes* (Bernstein 1971) the language use for meaning-making is mainly structured by the procedural rather than propositional

knowledge. Ways of knowing are bound by relational ideas and institutional routines, drawing upon communicative resources at higher cognitive levels. This contributes to understanding the relation between knowledge construction and discursive practice from a structural perspective, as suggested by Bernstein (1999) (Table 29):

Modality of discourse	Vertical discourse	Horizontal discourse
Distributive rules	Strong	Weak
Acquisition of knowledge	A graded performance through an integration of meanings based on specialized symbolic structures of knowledge	A general competence through the relevance of one contextual segment to another
Circulation of knowledge	Through explicit re-contextualization and evaluation	Through segmentally-related, context-specific and dependent orientation
Link of procedures	Hierarchical	Segmental
Social relation	Individual	Communicative

Table 29 Vertical discourse and horizontal discourse (Bernstein 1999)

The research findings suggest that the university small group talk in NUCASE is identified to have a hybrid discursive property in terms of knowledge construction: an integration of vertical discourse and horizontal discourse. Specifically, the talk is found to be mainly vertically structured with a discursively salient loading of procedures. Meanwhile, it shows features of horizontal discourse revealed from the patterned use of framing for locally organized and context-dependent meaning negotiation. It also should be noted that the hybrid discursive property does not necessarily see any internal boundary defined by academic disciplines.

Vertical discourse features in hierarchically linked procedures for explicit knowledge transmission and representation, with the acquisition achieved through "an integration of meanings based on specialized symbolic structures" (Bernstein 1999: 160). For example, Figure 25 shows how the two participants in Excerpt 6 explicate the procedure of working out the calculation results which are to be presented in a report. Joint attention is paid to the *transactional* aspect of the talk (Brown and Yule 1983; Lakoff 1989; see also McCarthy 2003) after the two participants have reached the point of minds-meeting from divergent framings concerning how much background information should be provided in their report. There is a linear order involved in the procedure which is discursively salient and traceable:

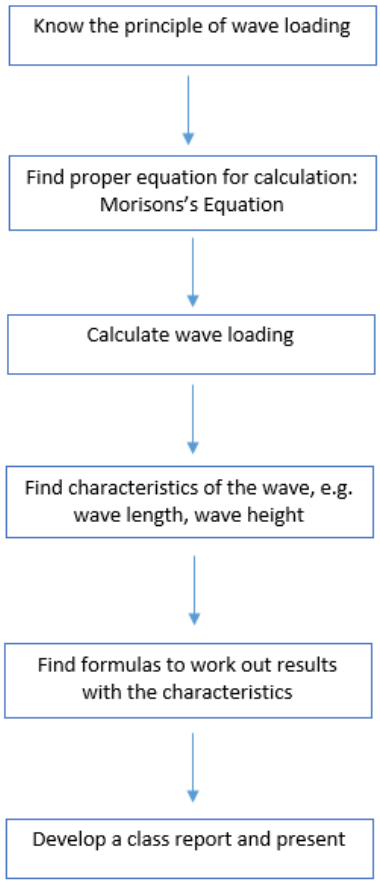


Figure 23 Vertical discourse in Excerpt 6 (BI)

The example above shows that the task procedure tightly binds the extended single turn of a certain speaker at a certain stage of the interaction, which constrains to a large extent other speakers' turn-taking opportunities (see Houtkoop and Mazeland 1985; Schegloff 1982). The form of knowledge requires a strong distributive rule of regulating the access and transmission of the knowledge (Bernstein 1999), which shapes the discourse into a hierarchical structure.

Following is another example (Excerpt 16, see Appendix D) of how vertical discourse is procedurally navigated through talk-in-interaction. In this example, three participants are working collaboratively to complete the task of designing a propeller. The basic design logic is explicated by \$1 at the start of the talk, which is "to design the propeller first and then choose the generator". Following this guiding principle, the other participants approach the procedure from distinct viewpoints. \$3 adopts a theoretical stance, trying to conceptualize the procedure from a bottom-up manner. \$6, on the other hand, takes a top-down perspective, focusing on the compatibility between different variables and actors in the modeling job at the empirical level. If \$3's relatively idealized portrayal of a linear path for the propeller design indicates a contextualization of general knowledge in the field into the locally distributive procedure, \$6's intention can be taken as a re-contextualization of the conceptualized procedures into a model for evaluation (see Bernstein 1999). One shall see that in the basic hierarchical structure, the talk undergoes a conceptualization circle concerning how professional knowledge is transmitted, evaluated and re-constructed in a particular task with real-world associations. Figure 26 visualizes the process.

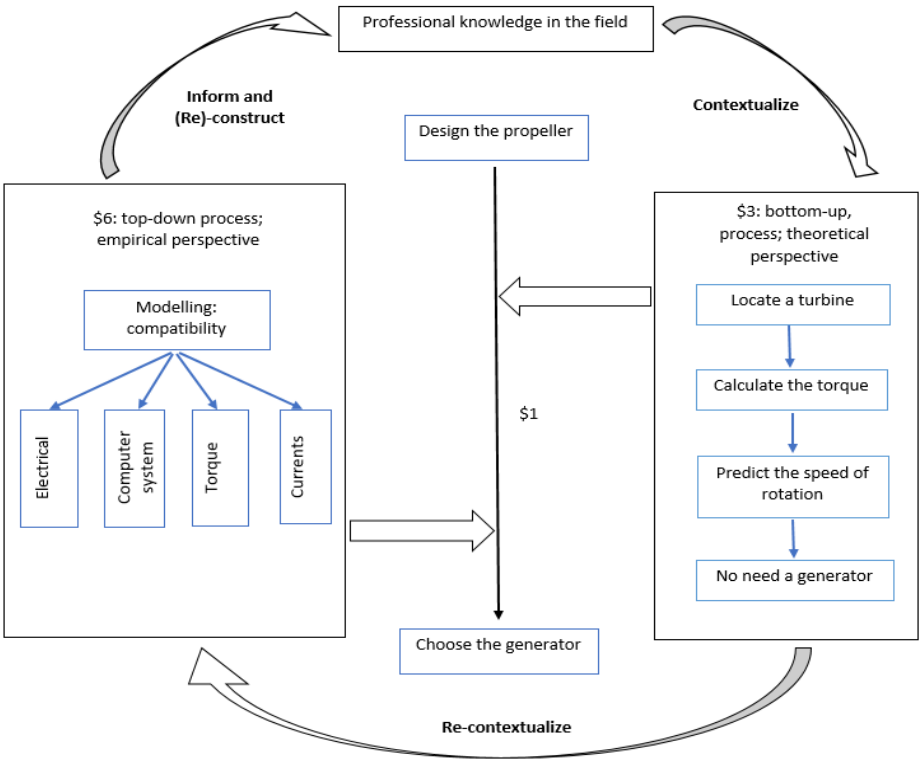


Figure 24 Vertical discourse in Excerpt 16 (ME)

Features of vertical discourse are also found in the applied-soft disciplines (see Becher 1987a). Excerpt 11 demonstrates how concepts are conveyed and explicated through the speaker's lexis use to construct knowledge in the monologic talk. In addition to strategic lexical choices and framing devices, it also demonstrates how forms of knowledge, specialized language use, and discourse patterns are associated with each other. Let us take a look at how different kinds of discursive elements are distributed and manipulated to shape the language for knowledge construction in this excerpt. According to Bernstein (1999), the integration of vertical discourses is "at the level of meanings" and "consists of specialized symbolic structures of explicit knowledge" through the "re-contextualization" of both the symbolic systems and the specialized modes of activity (p.161). Vertical discourses, to a large extent, are responsible for the target text's status of being "institutional"; or to paraphrase Bernstein (1999), "regulative" ("rules regulating access, transmission and evaluation", p.159). Table 30 shows clear evidence

of how the speaker shapes the talk structure in terms of vertical discursive elements deployment.

Criteria for a vertical discourse	**Examples in Excerpt 11 and explanation**
Coherence	Introducing *discourse markers* and *shared space* at the very beginning; defining *shared space (let me give a quick word about this; shared space means…)*; retrievable reference for textual cohesion (*one of the ways in which we do is we use words like…*); providing examples of discourse markers concerning how to create *shared space* (*you know, you see, I mean, kind of, etc.*); clarifying the viewpoint (*from a learning point of view…);* giving evaluation with emphasis (*really really important; very central*); emphasizing institutional association/implication (*…in academic er spoken discourse*); using discourse markers to signal discourse relations and functions (*and, so, yeah?,* etc.)
Explicitness	Explicate *shared space* as the core concept under discussion, referring directly to *learning* as an essential cognitive construal of *academic* properties
Systematic principled structure	N/A
Hierarchical organization	Scope of talk: *academic spoken discourse* Point of view and topic in the talk: *learning* Specific aspect (focus) of learning: *shared space* Mechanism of creation: *discourse markers*
Specialized language use	Using professional terms; providing definitions; explaining with examples; circular discourse organization
Specialized modes of interrogation	Goal-oriented information eliciting: persuasive guidance motivated by pedagogical purposes by strategically involving or developing a rapport with learners
Specialized criteria for text production and circulation	Explicit re-contextualization of *learning* as knowledge in terms of its potentially accessible and plastic spatial properties; explicit evaluation on creating *shared space* being *really really important* in learning, and on the functions of discourse markers being *very central* to involve learners

Table 30 Vertical discourse in Excerpt 11 (ALE)

Another observation is that the MARG talk in NUCASE shows curious characteristics in terms of its discursive structures. On the one hand, it reveals salient features of a vertical discourse: specialized language use, specialized modes for interrogation, specialized criteria for text production and evaluation. On the other hand, there is evidence of a horizontal discourse concerning the participants' contextualized orientation to and realization of the institutional talk, or as Arminen (2000) suggests, "CA studies use knowledge of the context anyway, either overtly or tacitly"(p. 436). The observation corresponds to Cosmay's (2005) finding that there is a linguistic dimension of "contextual, directive orientation vs. conceptual, informative focus" (p. 254) in university spoken discourse. The horizontal organization of the MARG talk falls towards the positive pole of the dimension, "associated with highly contextualized discourse" (p. 255) featuring the participants' active involvement, transfer of social resources/competence into the local context and negotiation of meaning based on shared community values. The vertical organization, by contrast, is found to be with negative loadings within the dimension, showing "a concentrated informational focus" (Cosmay 2005: 255) associated with "a preference for dense informational packaging" (p. 256; see also Biber 1988).

An intensive focus on the immediate sequential organizations of interaction in the audio data requires a conceptual structure which is linguistically represented by specialized language use of CA for interrogation and interpretation. However, the findings in this study seem to go against Cosmay's (2005) claim that the two poles of this textual dimension suggest "a fundamentally different approach to managing the concepts under discussion" (p. 256). From the perspective of framing, I argue that the "contextual, directive – conceptual, informative" dimension reveals the duality of discursive framing for conceptualization in context; namely,

cognitive structures applied and re-shaped in ongoing communicative actions. The two poles are thus not necessarily exclusive but interrelated and complementary.

A horizontal discourse is expected to entail rules of "the circulation of knowledge, behaviour and expectations" which are regulated according to the speaker's "status/position" concerning "social relations, practices and their contexts" (Bernstein 1999: 159). This proposition is essential to understand structures of institutional talk in terms of identity struggles (Benwell and Stokoe 2006) and power relations (Fairclough 1989; 1992). For example, the speaker in Excerpt 11 is the module leader, whose monologic talk, when examined in the whole small group session, overwhelmingly dominates the speaking time. The extended stretches of his talk provide a relatively large and flexible space, enabling the speaker to draw on as many informative resources and communicative techniques as he can to organize his teaching practice. A set of framing moves performed in such organization are clearly "local, context specific and dependent", with an intensive focus on "creating shared space of learning in academic spoken discourse" (Bernstein 1999: 159).

While the features are relatively representative, the competences of the speaker to convey and elaborate on the meaning of concepts are not readily generalizable. The reason is that these competences are "segmentally related" (Bernstein 1999: 160) rather than principled and integrated since patterns of pedagogical intervention are determined by different kinds of knowledge required to be constructed thus acquired by learners (Bernstein 1999). Such "segmental pedagogy" Bernstein (1999: 160) corresponds well to the dynamic and productive nature of the *classroom modes* (Walsh 2006) in terms of the role of language as the medium for knowledge transmission and construction. Therefore, interaction in any classroom as a discourse community can be viewed as horizontally mediated across various micro-contexts

(Stucky and Wimmer 2002) where "face-to-face meaning-making" contributes to "language socialization" for the co-construction of knowledge (Walsh and O'Keeffe 2007: 4).

Here is another example (Figure 27). The expanded talk based on Excerpt 14 demonstrates how knowledge is constructed in a contextually cyclic manner featuring segmented horizontal discursive elements. The discursive organization is perceived as horizontal because the knowledge circulated is through the relevance between contextualized discursive segments, rather than graded deployment of specific symbolic structures (see Bernstein 1999).

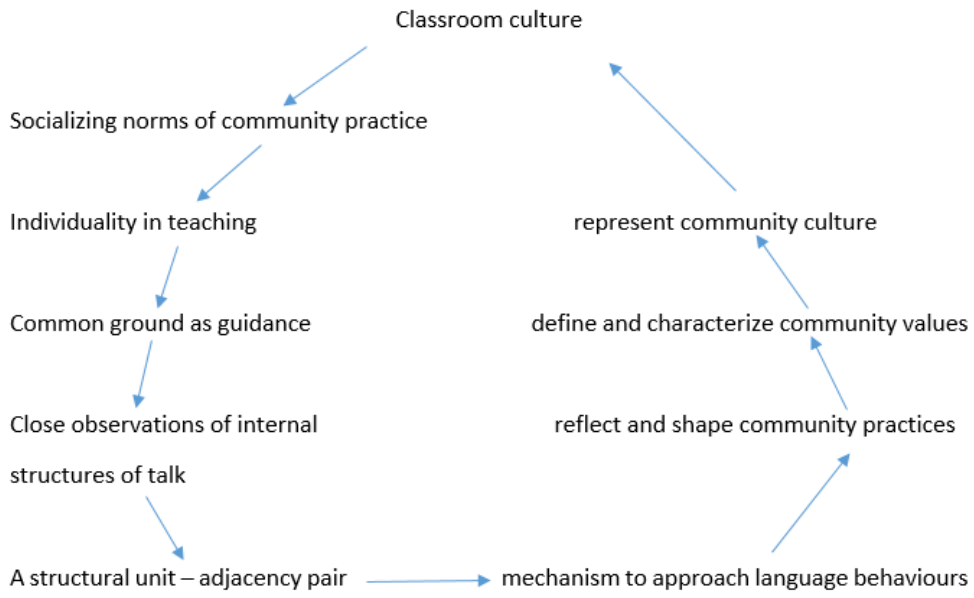


Figure 25 Horizontal discourse for knowledge construction in Excerpt 14 (ALE)

Features of horizontal discourse are found to be associated with the participants' framing-in-interaction, closely related to identities and speaker roles which are institutionally defined in the community practices. This corresponds to Bernstein's (1999) proposition that circulation of knowledge in horizontal discourse is basically realized through "segmental organization based on its contextual mode of orientation to social solidarity", based on "inter-dependence

between the social relationships" (p. 160-163). In this study, such segmental organization is found to break down the disciplinary barriers, revealing the participants' inter-contextually developed *interactional competence* (see Young 2000; 2011; 2013).

Framing orientations are closely associated with the participants' identities, especially when the speaker roles are defined prior to the talk. Duality of identity is represented in the community practice (Vygotsky 1978; Lave and Wenger 1991; Hyland 2002; Hartnell-Young 2003), revealed from the concern of sustaining one's "prominent position in cultural life" (Benwell and Stokoe 2002: 89; see also Gumperz 2003), based on the judgement of which aspect(s) of the identity is(are) relevant to certain discourse (see Linell 1998a, 1998b). The contrasting identities in this study are found to be more "productive" (Giddens 1981) than "oppressive" with "the transformative capacity of human action" (p. 67) to mutualize the procedural knowledge essential for successful communication. This proves the inter-dependence between social relationships and the forms of discourses (Bernstein 1999).

Different speaker roles usually imply sustaining different perspectives (MacLachlan and Reid 1994; Linell 1998b) around a single situation or concept, which is revealed from the participants' use of alternate framings in their talk. The attendance to the "multiple contexts" (see Goodwin and Goodwin 1992; Linnenbrink and Pintrich 2001) makes the progressive talk open to multiple interpretations within competing contextual spaces (Linell 1998a). The contrastive perspectives and interpretative mismatches (if there were any) offer the possibility of fresh perceptions (Bernstein 1999) which contribute to personalized meaning representations.

Development of procedural knowledge is facilitated by framing-in-interaction and achieved through the introduction of an altered way of representing speakership (Bernstein 1999). The university small group talk in NUCASE with tutor involvement is found to see a shifting interactional dynamics concerning the power distribution in Higher Education tutorial talk (Benwell and Stokoe 2006). Struggles are identified between the tutors' controlling of the agenda by preserving their traditionally defined roles (see Craig and Pitts 1990; Fassinger 1995; Basturkmen 1996; Benwell 1996) and the students' rising awareness and motivation of initiative-taking to manage *knowledge construction discourse* (van Aalst 2009) for individualized meaning representations.

The weak distributive rules of procedures involved in horizontal discourse (Bernstein 1999), on the other hand, may give an advantage to the more powerful social group in interaction. For example, when more than one tutor/delegation member is involved in the small group talk with students, co-framings are found to be more frequently and effectively used by the former. The tutors'/delegation members' co-framing behaviours may be deployed into divergent styles, drawing upon different aspects of the topic in focus, but are found to be converging at the conceptual level with a shared orientation to achieving particular "institutional agendas" (Heritage and Atkinson 1984: 15; Drew and Heritage 1992: 53; Kasper 2009: 15).

University small group talk, when analysed from the perspective of discourse structure, is found to be pervasively procedure-laden. The procedural aspect, however, is demonstrated in a dramatically different way from what has been identified in institutions below the Higher Educational level and in most language classrooms. According to Seedhouse (2004b), procedural talk in second language (L2) classroom contexts is aimed, predominantly from the teacher's side, to transmit information to the students concerning "the classroom activities

which are to be accomplished in the lesson" (p. 133). Similar features of procedural context have been found in Walsh's (2006) *managerial classroom mode* where the main pedagogical goal of information transmission is aligned to characteristic language use: "a single extended teacher turn" with "an absence of learner contributions" (p. 66).

6.3 Methodological considerations

In this section, some of the limitations of this study will be acknowledged. The shortcomings identified in the research design, however, are not presented to undermine the value of the study but rather to demonstrate that the use of research strategies would have an impact on the data interpretation.

First, the most obvious shortcoming of the study is that the data collection procedure was unable to capture everything happening in the small group talk when the participants were interacting. The absence of video data made it impossible to examine any embodied actions for meaning negotiation achieved in non-verbal ways. The audio recordings of the talk sessions were the only data available to the researcher, which rendered the analytical perspective limited while interaction is in nature multi-faceted. The lack of visual reproduction of the talk scenes may have interpretive consequences for the identification and characterization of discursive framings. For example, embodied actions such as head movements or gestures may indicate individual senses of spatial orientations which can project prepared actions to build upon or to shift a particular frame invoked by interlocutors. Similarly, video data are able to show the participants' actions when they are not talking thus can yield a significantly different analysis of intra/inter-turn silence. The silence attributable to either party not talking at a particular moment of interaction may be a result of the

participants' being busy doing something else (e.g. note taking, graph drawing, material seeking, etc.) thus does not necessarily indicate any trouble in communication.

Another matter that seemed problematic was the lack of knowledge regarding the specific roles of individual participants in particular tasks. While the participants involved in each small group talk session showed great homogeneity in terms of their age, academic background and language competence, the researcher was not able to know the exact labour divisions in a particular task especially for those talk sessions held on students' projects. As a result, examination of the participants' framing orientations and behaviours can only be conducted drawing upon roles which are naturally emerging from talk sequences. This would have an impact on the interpretation of the relationship between a speaker's role and what he/she says to frame the interaction.

Further considerations are given to the sampling procedure adopted in this study. To make the data analysis focused and the length of this thesis under control as required, this study only selected the over-used key semantic domains in NUCASE, namely, the semantic domains whose keyness values are larger than 0 when compared with the reference corpus. In addition, only the non-shared key semantic domains across the three sub corpora were selected for further analysis. Although the procedure has to a large extent ensured that the selected key semantic domains contain the most typical linguistic forms associated with lexical concepts which are highly discipline-specific, it did overlook the linguistic forms in the shared key semantic domains. The selection tendencies of the semantic potential of lexical concepts encoded by linguistic forms in the shared key semantic domains may be different from what has been found in the non-shared key semantic domains. The small group talk thus may be

partitioned in different ways with potential framing devices triggered by linguistic forms with salient psychological, emotional or social attributes rather than disciplinary features.

Finally, this study examined framing for knowledge construction through the participants' talk around discipline-specific lexical concepts. It should be noted that there are many ways in which frames can be invoked and shifted in talk-in-interaction and that making sense of particular lexical concepts is only one of them. Management of contextual resources can provide interlocutors with various indicators of frame activations and switches. Such contextual resources include but are not limited to code-switching to manifest speaker roles (Hancock 1997; Liebscher and Dailey-O'Cain 2003; Ellwood 2008), personal pronouns to shift individual and collective self-reference (Lerner and Kitzinger 2007), prosodic markers to distinguish activity types (Selting 1996; Günthner 1996; Uhmann 1996; Couper-Kuhlen 1996), repair sequences to manage mutual understanding (Seedhouse 2004a; Dings 2007; Seo 2008), and verbal art such as conversational joking (Rampton 1999; Bannink 2002; Davies 2003). More research is needed to address the multi-faceted nature of framing-in-interaction.

6.4 Recommendations for further research

In light of the above comments, a number of directions for further research are recommended. Firstly, it is strongly suggested that more research be conducted into university small group talk in terms of the relationship between knowledge construction and institutional routines. Being aware that the construction and dissemination of knowledge in university settings can be quite different from what has been practiced in institutions below Higher Education, researchers might consider conducting comparative studies to explore how different institutional cultures could possibly shape patterns of framing in teaching and learning activities. This would shed light on how learners at various levels can be encouraged and

facilitated to utilize linguistic and interactional resources for effective communication in classrooms or classroom-like settings. Noticing and processing cues of framing can help learners to mutualize expectations and negotiate meanings more efficiently, thus may contribute to a critical evaluation of pedagogy and active participation for collaborative learning.

As was suggested in the previous section, the key semantic domains with negative keyness values are worth investigating as well to explore why some linguistic forms with certain semantic senses are used significantly less frequently in the focus corpus when compared with the reference corpus. Meanwhile, the shared rather than non-shared key semantic domains across the academic disciplines may provide interesting findings too. The pilot study had identified 8 shared key semantic domains, among which the linguistic units in Domain A, E, Q, S, X, and Z (see Appendix A for details) may reveal distinct features of discursive framing, with the associative lexical concepts providing access sites to knowledge structures constituting concepts in linguistic, emotional, social and psychological realms. These lexical concepts are of great value to explore similarities across disciplinary discourses concerning how shared conceptual knowledge structures are taken into communicative contexts of institutional talk beyond the disciplinary boundaries. In addition, inquires of semantic frame application and framing-in-interaction based on key semantic domains needs further theoretical and empirical support, including the rationale of using frequency-based keyword procedures to explore the cognitive aspect of knowledge representation, construction, and transmission in social community practices.

Moving slightly outside the research setting of this study, comparative studies are necessary to gain a deeper understanding of the relationship between framings and knowledge

construction in discourse. For example, comparisons can be conducted between spoken discourse and written discourse, pure disciplines and applied ones, teacher talk and learner talk, native speakers and non-native speakers, institutional interactions and mundane conversations, professional experts and lay persons, etc. No matter what to compare, it is important for the researcher to clearly define the comparative parameters to ensure the validity of data analysis and interpretation. Additionally, a broader range of social variables which are external to the immediate sequences of interaction needs to be taken into account in such comparative studies. These social variables include but are not limited to gender, age, educational background, language competence, professional orientation, etc. As knowledge is socially constructed, these variables are important indicators of patterns of human reliance on structures for communication.

Finally, as suggested in the previous section, research on framing for knowledge construction beyond the lexical level is necessary. This involves taking into account other contextual resources that would have impacts upon framing patterns and discourse structures. As framing is closely associated with the pragmatics of interaction, it is equally important for the researcher to attend to the paralinguistic cues and multimodal aspects of talk in his/her investigation. This study has found a way-in to the data analysis using lexical concepts as discursive signposts, while the sequential analyses indeed revealed some features of multi-layered framing for meaning representation and negotiation. Research should continue to examine how interlocutors in collaborative talk draw on multiple resources to construct knowledge (Piirainen-Marsh and Tainio 2014). The integration of audio and video data is recommended to allow for a fuller representation of the multidimensionality of knowledge construction and the reality of interactional contexts (Balaman and Sert 2017), so that a better understanding of framing beyond the textual level can be achieved (see Knight 2009).

Chapter 7. Conclusion

In this final chapter, the aims of this study were revisited, and how these aims have been achieved will be discussed. In addition, the implications and contributions of the research findings were argued.

7.1 Research aims revisited

The principal objective of this study is to examine how the participants construct knowledge through the talk of discipline-specific lexical concepts in university small group interactions. This was regarded as a research endeavour to approach the participants' relative epistemic positions (Heritage 2013) at the interface between semantics research (e.g. Cruse 1986; Fillmore 1982a, 1982b; Barsalou 1992; Kittay and Lehrer 1992; Petruck 1996; Fauconnier and Turner 2006; Hamawand 2016) and the field of social interaction (e.g. Goffman 1974, 1986; Bateson 1972; Kendon 1992; Tannen 1993; Tannen and Wallat 1993; Gumperz 2003; Telles-Ribeiro and Hoyle 2009). The application and management of frames -or in other words, framing- plays a crucial role in linking the cognitive and communicative aspects of the processes of knowledge construction in the setting. Specifically, three research aims were outlined in Chapter 1:

- to explicate how specific lexical concepts provide access sites to particular cognitive frames so as to build models of disciplinary discourses;
- to examine how the participants invoke specific aspects of particular cognitive frames to represent meaning in a situated way; and
- to determine if, and how, the participants' management of sequences at talk would have impacts upon ways of framing-in-interaction for knowledge construction.

By investigating lexical concepts in the key semantic domains of each sub corpus, the first research aim was achieved. The findings suggest that meanings associated with linguistic forms (words and word clusters in this context) are mental representations of larger conceptual knowledge structures evoked by lexical concepts encoded by the linguistic forms thus perform a function of language in use. This helps to build models of disciplinary discourses in terms of the participants' lexical choices. Additionally, the analyses at this level demonstrate that semantic domains as coherent bodies of conceptual knowledge are useful to understand partitioning of discourse, therefore, contribute to identifying generic features of the latter.

In a further examination of the participants' lexical choices in the on-going discourse, the second research aim was achieved. This was about investigating how coherence was assigned to the discourse to assist local interpretation of individual meaning representations. The findings suggest that the interpretive frames invoked by the participants for situated meaning construction were a partial reflection of larger-scale conceptual knowledge structures. The semantic potential of lexical concepts provided access to the knowledge structures (frames) and determined what aspects of the structures (frame elements) were to be activated. The observations were considered in line with Coulson's (1997; 2006) argument that only a small portion of knowledge is available and accessible at a particular moment of on-line interpretation. One of the key findings from this study, in a further extension of this idea, is that building partial representations of cognitive frames is a defining feature of dynamic knowledge construction.

Finally, through close, detailed observations of the sequences of the participants' small group talk around target lexical concepts, the third research aim was achieved. The findings suggest

that the participants applied various framing devices to construct knowledge, including alternate framings of a single situation, co-framings, cross-domain mappings, mental space building and connection, and re-schematization of motivating circumstances. Specifically, the participants' sequential management of the small group talk showed particular relevance to shaping the patterns of alternative framings and co-framings. Analyses showed that, on the whole, the participants oriented to specific goals in their talk and took the construction of knowledge as a collaborative endeavour. Representation and negotiation of epistemic stances were found to be closely related to speaker roles and task nature, which demonstrated that framing for knowledge construction at the action level was institutionally defined and constrained.

7.2 Implications and contributions

This study has focused on the participants' discursive framing for knowledge construction in university small group talk. The investigations on the source, content, and transmission of knowledge contribute to broadening the scope of its characterization, concerning what would count as valid construction and dissemination of knowledge in educational practice (Peschl 2010). The research findings suggest that individualized meaning representations and meaning negotiation (Nunan 1988; Pica 1994; Holliday 1995; Long 1996; Fauconnier 1997) are essential in knowledge acquisition and development. This study also follows the call for a shift of the inquiry focus from products of pedagogy to processes of active participation for collaborative learning (Azmitia 1988; Resnick 1987; Ellis et al. 1993; Mercer 1995; Stahl 2004; Scardamalia and Bereiter 2006; Arvaja 2007; Cheung 2009; Sandra 2009; Roschelle and Teasley 2012). It takes into account the dynamics of small group interaction when exploring knowledge construction processes, which contributes to revealing the close associations between institutional procedures, speaker roles, and discourse structures.

The research findings have contributed to explaining the relationship between linguistically represented knowledge and the way language users conceptualize the world (Lucy 1997; Spuzic and Nouwens 2004; Wolff and Holmes 2011). At the centre of the relationship is the human reliance on structures to interpret real-world experiences, at both the cognitive level and the action level. Concepts take the form of conceptual structures which are semantically represented by lexical structures (Hamawand 2016). When embedded in sequential organizations of naturally occurring interaction, the semantic structures manifest pragmatic selectivity (Medin 1989; Jones and Smith 1993; Sweetser and Fauconnier 1996) drawing upon the *conditional relevance* for the analysis of the on-going production of social actions (Schegloff 1992).

This study has underlined the need to take discursive framing patterns into account to provide a more comprehensive structural description of knowledge in educational settings. Bernstein's (1999) categorization on discourse structures is based on the differential forms of knowledge, with the analytical perspective mainly at the macro level. This study, on the contrary, has particularly focused on the micro-mechanism of knowledge construction processes through the participants' framing-in-interaction. The analysis reveals that there are observable features found to be similar across the academic disciplines where discourse structures and knowledge forms mutually shape each other, and that different configurational elements of discourse are not necessarily associated with disciplinary cultures but more closely related to institutional routines and procedures to be practiced in goal-oriented tasks (Erickson and Shultz 1982; Drew and Heritage 1992; Benwell and Stokoe 2002; Seedhouse 2004b). In such sense, the "hard-soft" continuum to categorize disciplinary knowledge suggested by some scholars (e.g.

Bridgeman and Carlson 1984; Becher 1987; Casanave and Hubbard 1992; Hyland 1999) may need to be re-examined.

Additionally, this study has expanded the scope of traditional semantics research (e.g. Lexical Semantics, Cruse 1986) by integrating investigations on the applications of experience-based cognitive frames (Frame Semantics, Minsky 1975, 1977; Fillmore 1987a, 1987b; Barsalou 1992) evoked by linguistic forms and those on the approaches to discourse analysis through framing-in-interaction (Interactional Linguistics, Bateson 1972; Goffman 1986; Tannen 1993). A practice-based approach (Hanks 1996) to framing in discourse has been employed to achieve specific communicative agendas (Heller 2015). The use of Corpus Linguistics techniques has refreshed the received view of corpus data in traditional semantics research, manifesting that a corpus is by no means "dead language" (see Cruse 1986) but very much alive to provide rich linguistic evidence of patterned human cognitive and communicative behaviours. The Interactional Linguistics methods strive towards a more integrated approach to knowledge construction and meaning negotiation in talk-in-interaction. Detailed sequential analyses of the participants' small group talk contribute to revealing the interactional features of the genre as a variant of institutional talk.

This study has also broadened the empirical database of academic spoken interaction at the Higher Educational level. The observations correspond to theories of human social interaction for learning, adding to the claim that knowledge is socially constructed (Vygotksy 1978; Alexopoulou and Driver 1996; Bianchini 1997; Hyland 1999; Kelly and Crawford 1997; Kelly and Green 1998; Linn and Burbules 1993; Richmond and Striley 1996; Roth 1993; Mercer 1995; Kittleson and Southerland 2004; Lantolf and Thorne 2006). Inspired by the existing theories of nature of knowledge and learning, this study has gone further to highlight

the implicit (e.g. academic identity, discourse control) in addition to the explicit (e.g. rules and regulations, program development) institutional impacts on knowledge construction and exchange (see Goodson and Cole 1994; Antonek et al. 1997; Connelly and Clandinin 1999; Beijaard et al. 2000; Beijaard et al. 2004; O'Connor 2008; Eliot and Turns 2011).

Finally, this study has implications for understanding the communication and cooperation between different social institutions in terms of knowledge management. It is important to recognize that universities nowadays are deeply involved in the knowledge business (Laal 2011; Cranfield and Taylor 2008; see also Goddard 1998) and exposed to ever increasing marketplace pressures (Rowley 2000). Knowledge as an important intellectual capital and a meaningful economic resource has been taken as one of the most valuable assets of higher educational institutions as well as for other social organizations (Kidwell et al. 2000). The need to manage knowledge as an asset inevitably leads to heightened inter-institutional communication. In this study, such phenomena have been found in the small group talk sessions held between company delegations and university research teams. Distinct framing moves are identified to be associated with different institutional cultures, operational objectives, and speaker roles. The observation and interpretation of such framing behaviours are of great importance to understand in what way the knowledge possessed by one institutional group can be transmitted to another and how new knowledge can be co-constructed across different institutional groups. The engagement in knowledge management within and outside university settings, therefore, is expected to constitute an important dimension of Higher Education in the context of modern social interaction.

References

Abbott, A. (1988). *The System of the Professions: An Essay on the Division of Expert Labor*. Chicago, IL: University of Chicago Press.

Abdesslem, H. (1993). 'Analyzing foreign language lesson discourse'. *International Review of Applied Linguistics*, 31(3), 221-235.

Abelson, R. P. (1973). 'The structure of belief systems'. *Computer Models of Thought and Language*, 287-339.

Agirre, E. and Martinez, D. (2001). 'Learning class-to-class selectional preferences'. In Proceedings of CoNLL, Toulouse, France.

Al Hosni, S. S. R. (2014). 'EFL Learners' Negotiation of Meaning'. *International Journal of Applied Linguistics and English Literature*, 3(1), 215-223.

Alexopoulou, E. and Driver, R. (1996). 'Small-group discussion in physics: Peer interaction modes in pairs and fours'. *Journal of Research in Science Teaching*, 33, 1099–1114.

Allwood, J. (2003). 'Meaning potentials and context: Some consequences for the analysis of variation in meaning'. In H. Cuyckens, R. Dirven and J. Taylor (eds.). *Cognitive Approaches to Lexical Semantics*. Berlin: Mouton de Gruyter, 29-66.

Altenberg B. and Granger S. (2002) 'Recent trends in cross-linguistic lexical studies'. In Altenberg B. and Granger S. (eds) *Lexis in Contrast. Corpus-based Approaches. Studies in Corpus Linguistics 7*. Amsterdam: Benjamins. 3-48.

Amato, C. H. and Amato, L. H. (2005). Enhancing student team effectiveness: application of Myers-Briggs personality assessment in business courses. *Journal of Marketing Education*, 27(1), 41-51.

Andreevskaia, A. and Bergler, S., (2007). 'CLaC and CLaC-NB: Knowledge-based and corpus-based approaches to sentiment tagging'. In *Proceedings of the 4th international workshop on semantic evaluations* (pp. 117-120). Association for Computational Linguistics.

Andrew, D.P.S., Pedersen, P.M. and McEvoy, C.D. (2011). *Research Methods and Design in Sport Management*. Human Kinetics.

Anton, M. (1999). 'The discourse of a learner-centered classroom: Sociocultural perspectives on teacher-learner interaction in the second-language classroom'. *Modern Language Journal*, 303-318.

Antonek, J. L., McCormick, D.E. and Donato, R. (1997). 'The student teacher portfolio as autobiography: Developing a professional identity'. *The Modern Language Journal*, 81(1), 15-27.

Arminen, I. (2000). 'On the context sensitivity of institutional interaction'. *Discourse & Society*, 11(4), 435-458.

Armstrong, S. L., Gleitman, L. R. and Gleitman, H. (1983). 'What some concepts might not be'. *Cognition*, 13, 263-308.

Arvaja, M. (2007). 'Contextual perspective in analysing collaborative knowledge construction of two small groups in web-based discussion'. *International Journal of Computer-Supported Collaborative Learning*, 2(2-3), 133-158.

Arvaja, M., Salovaara, H., Hakkinen, P. and Jarvela, S. (2007). 'Combining individual and group-level perspectives for studying collaborative knowledge construction in context'. *Learning and Instruction*, 17, 448-459.

Asmuth, J.A. and Gentner, D. (2005). 'Context sensitivity of relational nouns'. In *Proceedings of the Cognitive Science Society*, 163-168.

Atkins, S., Rundell, M., and Sato, H. (2003).' The contribution of FrameNet to practical lexicography'. *International Journal of Lexicography*, 16(3), 333-357.

Atkinson, J. M. and Heritage, J. C. (eds.) (1984). *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press.

Atwood, S., Turnbull, W. and Carpendale, J. I. (2010). 'The construction of knowledge in classroom talk'. *The Journal of the Learning Sciences*, 19(3), 358-402.

Azmitia, M. (1988). 'Peer interaction and problem solving: When are two heads better than one?'. *Child Development*, 59, 87-96.

Baker, C., Ellsworth, M. and Erk, K., (2007). 'SemEval'07 task 19: frame semantic structure extraction'. In *Proceedings of the 4th International Workshop on Semantic Evaluations* (pp. 99-104). Association for Computational Linguistics.

Baker, P. (2004). *Public Discourses of Gay Men*. London: Routledge.

Baker, P. (2006). *Using Corpora in Discourse Analysis*. A&C Black.

Baker, P. (2015). Personal communication at Lancaster Summer School for Corpus Linguistics Approaches to Social Sciences, Lancaster University, UK.

Baker, P. (2016). 'The shapes of collocation'. *International Journal of Corpus Linguistics*, 21(2), 139-164.

Baker, P., Gabrielatos, C., Khosravinik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refuggees and asylum seekers in the UK press. *Discourse and Society*, 19(3), 273 – 306.

Baker, P., Hardie, A. and McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Balaman, U. and Sert, O. (2017). 'The coordination of online L2 interaction and orientations to task interface for epistemic progression'. *Journal of Pragmatics*, http://dx.doi.org/10.1016/j.pragma.2017.01.015

Baldry, A. (2000). *Multimodality and Multimediality in the Distance Learning Age*. Campobasso: Palladino.

Bannink, A. (2002). 'Negotiating the paradoxes of spontaneous talk in advanced L2 classes'. In Kramsch (ed.), *Language Acquisition and Language Socialization*. London: Continuum. 266-268.

Barron, B. (2003). When smart groups fail. *The Journal of Learning Science*, 12(3), 307-359.

Barsalou, L. W (1987). 'The instability of graded structure: Implications for the nature of concepts'. In U. Neisser (ed), *Concepts and Conceptual Development*. New York: Cambridge University Press. 101-140.

Barsalou, L. W. (1982). 'Context-independent and context-dependent information in concepts'. *Memory and Cognition*, 10, 82-93.

Barsalou, L. W. (1992). 'Frames, concepts and conceptual fields'. In Lehrer, A. and Kittay, E. F. (eds). *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organizations*. 21-74.

Barsalou, L.W. (1989). 'lntra-coneept similarity and its implications for inter-concept similarity'. In Vosniadou, S. and Ortony, A. (eds.), *Similarity and Analogical Reasoning*. New York: Cambridge University Press. 76-121.

Barth-Weingarten, D. (2008). 'Interactional Linguistics'. In Antos, G., Ventola, E. and Weber, T. (eds.), *Handbook of interpersonal communication*. Berlin: Walter de Gruyter. 77-106.

Basturkman, H. (2002). "Negotiating meaning in seminar-type discussion and EAP". *English for Specific Purposes*, 21, 233–242.

Basturkmen, H. (1996). 'The Discourse of Academic Seminars: Structures and Strategies of Interaction'. Unpublished PhD thesis, Aston University, Birmingham.

Bateson, G. (1972). *Steps to an Ecology of Mind*. New York: Ballantine Books.

Beach, W. (1993). 'Transitional regularities for 'casual' "Okay" usages'. *Journal of Pragmatics*, 19, 325–352.

Beach, W. and Metzger, T. (1997). Claiming insufficient knowledge. *Human Communication Research*, 23 (4), 562-588.

Beamer, B., Bhat, S., Chee, B., Fister, A., Rozovskaya, A. and Girju, R., (2007). 'UIUC: A knowledge-rich approach to identifying semantic relations between nominals'. In *Proceedings of the 4th International Workshop on Semantic Evaluations* (pp. 386-389). Association for Computational Linguistics.

Becher, T. (1987a). 'The disciplinary shaping of the profession'. In B. R. Clark (ed.) *The Academic Profession*. Berkeley, CA: University of California Press. 271-303.

Becher, T. (1987b). 'Disciplinary discourse'. *Studies in Higher Education*, 12(3), 261-274.

Becher, T. (1989). *Academic Tribes and Territories*. Milton Keynes: SRHE/Open University Press.

Becher, T. (1994). 'The significance of disciplinary differences'. *Studies in Higher Education*, 19(2), 151-161.

Becher, T. and Trowler, P. (2001). *Academic Tribes and Territories: Intellectual Enquiry and the Culture of Disciplines*. McGraw-Hill Education (UK).

Bednarek, M. (2005). 'Frames revisited – the coherence-inducing function of frames'. *Journal of Pragmatics*, 37, 685-705.

Bednarek, M. (2008). 'Semantic preference and semantic prosody re-examined'. *Corpus Linguistics and Linguistics Theory*, 4(2), 119-139.

Behfar, K., Kern, M. and Brett, J. (2006). "Managing challenges in multicultural teams". In Y. Chen (ed.), *Research on managing groups and teams: National culture and groups*. 239–269.

Beijaard, D., Meijer, P.C. and Verloop, N. (2004). 'Reconsidering research on teachers' professional identity'. *Teaching and Teacher Education*, 20(2), 107-128.

Beijaard, D., Verloop, N. and Vermunt, J.D. (2000). 'Teachers' perceptions of professional identity: An exploratory study from a personal knowledge perspective'. *Teaching and Teacher Education*, 16(7), 749-764.

Benwell, B. (1996). The Discourse of University Tutorials. Unpublished PhD thesis, University of Nottingham, UK.

Benwell, B. and Stokoe, E. (2002). 'Constructing discussion tasks in university tutorials: Shifting dynamics and identities'. *Discourse Studies*, 4, 429-453.

Bereiter, C. (2002). 'Design research for sustained innovation'. *Cognitive Studies*, 9(3), 321-327.

Bereiter, C. and Scardamalia, M. (1989). 'Intentional learning as a goal of instruction'. *Knowing, learning, and instruction: Essays in honor of Robert Glaser*, 361-392.

Bereiter, C. and Scardamalia, M. (1996). 'Rethinking learning'. *The Handbook of Education and Human Development: New Models of Learning, Teaching and Schooling*, 1, 485-513.

Berge, Z. (1999). "Interaction in post-secondary web-based learning". *Education Technology*, 39(1), 5–10.

Bergen, B. K. and Chang, N. C. (2005). ‘Embodied construction grammar in simulation-based language understanding’. In J.-O-O ¨stman and M. Fried (eds.), *Construction Grammars*. Amsterdam: Benjamins, 147–190.

Berger, P. L. and Luckmann, T. (1967). *The Social Construction of Reality*. London: Allen Lane.

Berkenkotter, C. and Huckin, T. N. (1995). *Genre Knowledge in Disciplinary Communication: Cognition/culture/power*. Lawrence Erlbaum Associates, Inc.

Bernstein, B. (1971). ‘On the classification and framing of educational knowledge’. In Young, M. F. (ed.) *Knowledge and Control: New Directions for the Sociology of Education*. London: Collier-Macmillan. 47-69.

Bernstein, B. (1999). ‘Vertical and horizontal discourse: An essay’. *British Journal of Sociology of Education*, 20(2), 157-173.

Berrill, D. P. (1991). ‘Exploring underlying assumptions: Small group work of university undergraduates’. *Educational Review*, 43(2), 143-157.

Bianchini, J. (1997). ‘Where knowledge construction, equity, and context intersect: Student learning of science in small groups’. *Journal of Research in Science Teaching,* 43, 1039–1065.

Biber, D. (1993). ‘Co-occurrence patterns among collocations: A tool for corpus-based lexical knowledge acquisition’. *Computational Linguistics*, 19(3), 531-538.

Biber, D. (2009). ‘A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing.’ *International Journal of Corpus Linguistics,* 14(3), 275-311.

Biber, D. (2009). ‘A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing’. *International Journal of Corpus Linguistics*. 14(3). 275-311.

Biber, D., Conrad, S., and Reppen, R. (1996). Corpus-based investigations of language use. *Annual Review of Applied Linguistics,* 16, 115-136.

Biber, D., Conrad, S., and Reppen, R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press, Cambridge.

Biggs, J. B. (1987). 'Student Approaches to Learning and Studying'. Research Monograph. Australian Council for Educational Research Ltd., Radford House, Frederick St., Hawthorn 3122, Australia.

Blakemore, D. (1996). 'Are apposition markers discourse markers?'. *Journal of linguistics*, *32*(2), 325-347.

Blakemore, D. (2000). 'Indicators and procedures: nevertheless and but'. *Journal of Linguistics,* 36(3), 463-486.

Blakemore, D. (2002). *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.

Blatchford, P., Kutnick, P., Baines, E. and Galton, M. (2003). 'Toward a social pedagogy of classroom group work'. *International Journal of Educational Research,* 39(1), 153-172.

Bocchi, G. and Cianci, E. (2012). 'Introduction: Contexts, Boundaries, and Knowledge Construction'. *World Futures,* 68(3), 145-158.

Boekaerts, M. and Minnaert, A. (2006). Affective and motivational outcomes of working in collaborative groups. *Educational Psychology*, 26(2), 187-208.

Bolinger, D. (1957). *Interrogative Structures of American English*. Tuscaloosa, AL: University of Alabama Press.

Bondi, M. (2010). 'Perspectives on keywords and keyness'. In Bondi, M. and Scott, M. (eds.) *Keyness in Texts*. John Benjamins. 1-18.

Borsley, R. D. and Ingham, R. (2002). ‘Grow your own linguistics? On some applied linguistics views of the subject’. *Lingua Franca*, 112, 1-6.

Boud, D., Cohen, R. and Sampson, J. (1999). Peer learning and assessment. Assessment and *Evaluation in Higher Education*, 24(4), 413-426.

Bowker, G. and Star, S. L. (1999). *Sorting Things Out: Classification and its Consequences*. Cambridge: The MIT Press.

Brandt, A. (2011). ‘The maintenance of mutual understanding in online second language talk’. Unpublished PhD thesis, Newcastle University, UK.

Brewer, M.B. and Gardner, W. (1996). ‘Who is this" We"? Levels of collective identity and self-representations’. *Journal of Personality and Social Psychology*, 71(1), 83-93.

Brezina, V., McEnery, T. and Wattam, S. (2015). ‘Collocations in context: A new perspective on collocation networks’. *International Journal of Corpus Linguistics*, 20(2), 139-173.

Bridgeman, B. and Carlson, S. (1984). ‘Survey of academic writing tasks’. *Written Communication*, 1(2), 247-280.

Brinton, L. J. (2000). *The structure of modern English: a linguistic introduction*. Illustrated edition. John Benjamins Publishing Company.

Brophy, J. E., and Evertson, C. M. (1978). ‘Context variables in teaching’. *Educational Psychologist*, 12(3), 310–16.

Brown, G. And Yule, G. (1983). *Discourse Analysis*. Cambridge: Cambridge University Press.

Brown, P. and Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage* (Vol. 4). Cambridge: Cambridge University Press.

Bublitz,W. (1995). ‘Semantic prosody and cohesive company: somewhat predictable’.

General and Theoretical Papers 347. Duisburg: L.A.U.D. (Linguistic Agency University of Duisburg), 1–23.

Burr, V. (1995). *An Introduction to Social Constructionism*. London: Routledge.

Burton-Roberts, N. (1975). 'Nominal apposition'. *Foundations of Language*,13(3), 391-419.

Button, G. and Casey, N. (1988). "Topic initiation: business-at-hand". *Research on Language and Social Interaction*, 22, 61–92.

Cameron, L. and Deignan, A. (2006). 'The emergence of metaphor in discourse'. *Applied Linguistics*, 27(4): 671-690.

Cantwell, R. H. and Andrews, B. (2002). Cognitive and psychological factors underlying secondary students' feelings towards group work. *Educational Psychology*, 22(1), 75-91.

Carragee, K.M. and Roefs, W. (2004). 'The neglect of power in recent framing research'. *Journal of communication*, 54(2), 214-233.

Casanave, C. and Hubbard, P. (1992). 'The writing assignments and writing problems of doctoral students: Faculty perceptions, pedagogical issues and needed research'. *English for Specific Purposes*, 11(1), 33-49.

Cazden, C. (1986). 'Classroom discourse'. In M. C. Wittrock (ed.), *Handbook of Research on Teaching* (3rd ed.). NY: MacMillan. 432-463.

César, M. and Santos, N. (2006). 'From exclusion to inclusion: Collaborative work contributions to more inclusive learning settings'. *European Journal of Psychology of Education*, 21(3), 333-346.

Chang, B. (2010). 'Culture as a tool: facilitating knowledge construction in the context of a learning community'. *International Journal of Lifelong Education*, 29(6), 705-722.

Chang-Wells, G. L. M. and Wells, G. (1993). 'Dynamics of discourse: Literacy and the construction of knowledge'. In Forman, E. A., Minick, N. and Stone, C. A. (eds.) *Contexts for*

Learning: Sociocultural Dynamics in Children's Development. New York: Oxford University Press. 58-90.

Channell, J. (2000). 'Corpus-based analysis of evaluative lexis'. In Hunston, S. and G. Thompson (eds.) *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press, 38–55.

Charles, M. (2006). 'Phraseological patterns in reporting clauses used in citation: A corpus-based study of theses in two disciplines'. *English for Specific Purposes*, 25(3), 310-331.

Cheung, M. (2009). Merging Corpus Linguistics and collaborative knowledge construction. Unpublished PhD thesis, the University of Birmingham, UK.

Christodoulidou, M. (2011). 'Lexical markers within the university lecture'. *Novitas- ROYAL (Research on Youth and Language),* 5 (1), 143-160.

Cienki, A. (2012). 'Frames, Idealized Cognitive Modes and Domains'. In Geeraerts, D. and Cuyckens, H. (eds.) *The Oxford Handbook of Cognitive Linguistics*, online version.

Clair, R. P. (1993). 'The use of framing devices to sequester organizational narratives: Hegemony and harassment'. *Communication Monographs*, 60(2), 113-136.

Clark, H. H.and Brennan, S. E. (1991). 'Grounding in communication'. In L. B. Resnick, J. M. Levine, & S. D. Teasley (eds.), *Perspectives on Socially Shared Cognition*. Washington, D. C.: American Psychological Association. 127-149.

Clark, M. (2003). 'Computer science: A hard-applied discipline?'. *Teaching in Higher Education*, 8(1), 71-87.

Clarks, H. (1983). 'Making sense of nonce sense'. In Levelt, W. J. M. and R. Jarvella (eds.), *The Process of Language Understanding*. New York: Wiley, 297–331.

Clausner, T. C. and Croft, W. (1999). 'Domains and image schemas'. *Cognitive Linguistics*, 10, 1-32.

Cobb, P. and Yackel, E. (1996). 'Constructivist, emergent, and sociocultural perspectives in the context of developmental research'. *Educational Psychologist,* 3(4), 175–190.

Cohen, E. (1994). *Designing Group Work* (2nd ed.). New York: Teachers College Press.

Cole, M. (1996). *Cultural Psychology: A Once and Future Discipline*. Cambridge MA: Harvard.

Connelly, F. M. and Clandinin, D. J. (1999). *Shaping a Professional Identity: Stories of Educational Practice*. Teachers College Press.

Considine, D. (1994). 'The Loose Cannon Syndrome: University as Business and Students as Consumers', *Australian Universities' Review*, 37(1), 36–40.

Cook, G. (2009). *Discourse*. Oxford: Oxford University Press.

Coulson, S. (1997). 'Semantic leaps: The rule of frame-shifting and conceptual blending in meaning construction'. Unpublished PhD thesis, University of California, San Diego.

Coulson, S. (2001). *Semantic Leaps: Frame-shifting and Conceptual Blending in Meaning Construction*. Cambridge: Cambridge University Press.

Coulson, S. (2006). 'Constructing meaning'. *Metaphor and Symbol*, 21(4), 245-266.

Coultard, M. (1992). *Advances in Spoken Discourse Analysis.* London: Routledge.

Couper-Kuhlen, E. (1996).The prosody of repetition: On quoting and mimicry. In E. Couper-Kuhlen & M. Selting (eds.), *Prosody in Conversation: Interactional studies*. Cambridge: Cambridge University Press. 366-405.

Cowles, H.W., Walenski, M. and Kluender, R. (2007). 'Linguistic and cognitive prominence in anaphor resolution: Topic, contrastive focus and pronouns'. *Topoi*, *26*(1), 3-18.

Craig, D. and Pitts, M.K. (1990). 'The Dynamics of Dominance in Tutorial Discussions', *Linguistics,* 28, 125–138.

Cranfield, D. J. and Taylor, J. (2008). 'Knowledge Management and Higher Education: A UK Case Study'. *Journal of Knowledge Management*, 6(2), 85-100.

Croft, W. (1993). 'The role of domains in the interpretation of metaphors and metonymies'. *Cognitive Linguistics*, 4, 335–370.

Croft, W. (2000). *Explaining Language Change: An Evolutionary Approach*. London: Longman.

Cruse, D. A. (1986). *Lexical Semantics*. Cambridge University Press.

Csomay, E. (2005). 'Linguistic variation within university classroom talk: A corpus-based perspective'. *Linguistics and Education*, 15, 243-274.

Cuadros, M. and Rigau, G. (2007) 'Semeval-2007 task 16: Evaluation of wide coverage knowledge resources'. In *Proceedings of the 4th International Workshop on Semantic Evaluations* (pp. 81-86). Association for Computational Linguistics.

Cuadros, M., Padr´o, L. and Rigau, G. (2005). 'Comparing methods for automatic acquisition of topic signatures'. In Proceedings of RANLP, Borovets, Bulgaria.

Cutting, J. (2001). 'The speech acts of the in-group'. *Journal of Pragmatics*, 33(8), 1207-1233.

Cutting, J. (2015). 'Dingsbums und so: Beliefs about German vague language'. *Journal of Pragmatics*, 85, 108-121.

Davies, C. E. (2003). 'How English-learners joke with native speakers: an interactional sociolinguistic perspective on humor as collaborative discourse across cultures'. *Journal of Pragmatics,* 35, 1361-1385.

de Bruin, J. and Scha, R. (1988). 'The interpretation of relational nouns'. In *Proceedings of the 26th annual meeting on Association for Computational Linguistics* (pp. 25-32). Association for Computational Linguistics.

DeKlerk, V. (1995). “The discourse of postgraduate seminars”. *Linguistics and Education*, 7, 157–174.

Denzin, N. K. (1988). ‘Qualitative analysis for social scientists’. *Contemporary Sociology*, 17(3), 430-432.

Denzin, N. K. and Keller, C. M. (1981). ‘Frame analysis reconsidered’. *Contemporary Sociology*, 10(1), 52-60.

Deppermann, A. (2005). ‘Conversational interpretation of lexical items and conversational contrasting’. In Hakulinen, A. and Selting, M. (eds.) *Syntax and Lexis in Conversation*. Amsterdam: Benjamins. 289-317.

Dillenbourg, P., Baker, M., Blaye, A. and O’Malley, C. (1996). The evolution of research on collaborative learning. In P. Reimann, & H. Spada (eds.), *Learning in Humans and Machines*. London: Pergamon. 189-211.

Dings, A. (2007). ‘Developing interactional competence in a second language: A case study of a Spanish language learner’. Unpublished doctoral dissertation, University of Texas, Austin.

Dole, J. A. and Sinatra, G. M. (1998). ‘Reconceptalizing change in the cognitive construction of knowledge’. *Educational Psychologist*, 33(2-3), 109-128.

Donnellan, K. S. (1977). ‘Reference and definite descriptions’. In S. P. Schwartz (ed.), *Naming, Necessity, and Natural Kinds*. Ithaca, NY: Cornell University Press. 42-65.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies*. Oxford: Oxford University Press.

Dörnyei, Z. and Murphey, T. (2003). *Group Dynamics in the Language Classroom*. Ernst Klett Sprachen.

Douglas, M. T. (1986). *How Institutions Think*. Syracuse, NY: Syracuse University Press.

Drew, P. (1991). ‘Asymmetries of knowledge in conversational interactions’. In I.Markova & K. Foppa (eds.), *Asymmetries in dialogue*. Hemel Hempstead, England: Harvester/Wheatsheaf. 29-48.

Drew, P. and Heritage, J. (1992). *Talk at Work: Interaction in Institutional Settings*. Cambridge England; New York: Cambridge University Press.

Duff, P. (2006). ‘Beyond generalizability: contextuliazation, complexity, and credibility in applied linguistics research’. In Chalhoub-Deville, M., Chapelle, C. A., Duff, P. (eds). *Inference and Generalizability in Applied Linguistics: Multiple Perspectives*. Amsterdam: John Benjamins.

Duff, P. A. (2002). ‘The discursive co-construction of knowledge, identity, and difference: An ethnography of communication in the high school mainstream’. *Applied Linguistics,* 23(3), 289-322.

Duranti, A. (2006). ‘Transcripts, like shadows on a wall’. *Mind, Culture, and Activity, 13*(4), 301-310.

Ebenezer, J., Chacko, S., Kaya, O.N., Koya, S.K. and Ebenezer, D.L. (2010). ‘The effects of common knowledge construction model sequence of lessons on science achievement and relational conceptual change’. J*ournal of Research in Science Teaching,* 47(1), 25-46.

Edwards, A. and Westgate, D. (1994). *Investigating Classroom Talk*. London: Falmer.

Eliot, M. and Turns, J. (2011). ‘Constructing professional portfolios: Sense-making and professional identity development for engineering undergraduates’. *Journal of Engineering Education*, 100(4): 630-654.

Ellis, R. (2003). *Task-based Language Teaching and Learning*. Oxford: Clarendon Press.

Ellis, S., Klahr, D. and Siegler, R. S. (1993). ‘Effects of feedback and collaboration on changes in children's use of mathematical rules’. A paper presented in Society for Research in Child Development. New Orleans.

Ellwood, C. (2008). Questions of classroom identity: what can be learned from code-switching in classroom peer group talk? *The Modern Language Journal*, 92, 538-557.

Engestrom, Y. (1999). ‘Activity theory and individual and social transformation’. In Y. Engstrom, R. Miettinen and R. Punamaki (eds.) *Perspectives on Activity Theory*. NY: Cambridge University Press. 19-38.

Erickson, F. and Shultz, J. (1982). *The Counselor as Gatekeeper: Social Interaction in Interviews*. New York: Academic Press.

Evans, V. (2004a). *The Structure of Time: Language, Meaning and Temporal Cognition*. Amsterdam: John Benjamins.

Evans, V. (2004b).’ How we conceptualise time’. *Essays in Arts and Sciences*, 33(2), 13–44.

Evans, V. (2005). ‘The meaning of time: Polysemy, the lexicon and conceptual structure’. *Journal of Linguistics*, 41(1), 1–39.

Evans, V. (2006). ‘Lexical concepts, cognitive models, and meaning construction’. Cognitive Linguistics, 17(4), 491-534.

Evans, V. (2009). *How Words Mean: Lexical Concepts, Cognitive Models, and Meaning Construction*. Oxford: Oxford University Press.

Evans, V. and Green, M. (2006). *Cognitive Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

Evert, S. (2010). Computational Approaches to Collocations. Retrieved from http://www.collocations.de/ (last accessed March 2015).

Evison, J. (2012). ‘A corpus linguistic analysis of turn-openings in spoken academic discourse: understanding discursive specialisation’. *English Profile Journal*. 3(1), e4.

Evison, J. (2013). ‘Turn openings in academic talk: where goals and roles intersect’. *Classroom Discourse*. 4(1), 3-26.

Evnitskaya, N. and Morton, T. (2011). ‘Knowledge construction, meaning-making and interaction in CLIL science classroom communities of practice’. *Language and Education*, 25(2), 109-127.

Fairclough, N. (1989). *Language and Power*. London and New York: Longman.

Fairclough, N. (1992). *Critical Language Awareness*. New York: Longman.

Fassinger, P. A. (1995). ‘Understanding Classroom Interaction: Students’ and Professors’ Contributions to Students’ Silence’, *Journal of Higher Education*, 66(1), 82-96.

Fauconnier, G. (1992). ‘Sens potential: Grammaire et discours’. In De Mulder, Schuerewegen, F. and Tasmowski, L. (eds.). *Enunciation et Parti Pris*. Amsterdam: Rodopi.

Fauconnier, G. (1994). *Mental Spaces: Aspects of Meaning Construction in Natural Language*. Cambridge University Press.

Fauconnier, G. (1997a). ‘Mappings’. In Fauconnier, G. *Mappings in Thought and Language*. Cambridge: Cambridge University Press. 1-33.

Fauconnier, G. (1997b). ‘Mental-space connections’. In Fauconnir, G. *Mappings in Thought and Language*. Cambridge: Cambridge University Press. 34-71.

Fauconnier, G. (2012). ‘Mental Spaces’. In Geeraerts, D. And Cuyckens, H. (eds.) *The Oxford Handbook of Cognitive Linguistics*, online version.

Fauconnier, G. and Turner, M. (2006). ‘Mental Spaces’. In Geeraerts, D., Dirren, R. and Taylor, J. R. (eds.) *Cognitive Linguistics: Basic Readings*. Berlin; New York: Monton de Gruyter. 303-371.

Fillmore, C. J. (1961). *Indirect Object Constructions in English and the Ordering of Transformations*. The Hauge: Mouton.

Fillmore, C. J. (1963). 'The position of embedding transformations in a grammar'. *Word,* 19, 208-301.

Fillmore, C. J. (1968). 'The case for case'. In Bach, E. and Harms, R. (eds.) *Universals in Linguistic Theory*. New York: Holt Rinehart and Winston. 1-88.

Fillmore, C. J. (1969/1971). 'Verbs of judging: An exercise in semantic description'. In Charles J. Fillmore and D. Terrence Langendoen (eds.) *Studies in Linguistic Semantics*. New York: Holt, Rinehart and Winston. 272–289.

Fillmore, C. J. (1975). An alternative to checklist theories of meaning. Proceedings of the First Annual Meeting of the Berkeley Linguistics Society, ed. by Cathy Cogen et al., 123-131. Berkeley: Berkeley Linguistics Society.

Fillmore, C. J. (1977). "Scenes-and-frames semantics". In: A. Zampolli, ed., *Linguistic Structures Processing*, 55-81. Amsterdam: North-Holland.

Fillmore, C. J. (1978). On the organization of semantic information in the lexicon. In Papers from the Parasession on the Lexicon, 1–11. Chicago: The Chicago Linguistic Society.

Fillmore, C. J. (1982a). 'Frame Semantics'. In Geeraerts, D., Dirren, R. and Taylor, J. R. (eds.) *Cognitive Linguistics: Basic Readings*. Berlin; New York: Monton de Gruyter. 373-400.

Fillmore, C. J. (1982b). 'Frame semantics'. In *Linguistics in the Morning Calm*, ed. by The Linguistic Society of Korea. Soeul: Hanshin. 111-137.

Fillmore, C. J. (1985a). 'Frames and the semantics of understanding'. *Quaderni di Semantica* 6(2), 222-254.

Fillmore, C. J. (1985b). ‘Syntactic intrusions and the notion of grammatical construction’. In proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society, ed. by Mary Niepokuj, et. al., 73-86. Berkeley: Berkeley Linguistics Society.

Fillmore, C. J. (1986). ‘U-semantics, Second Round’. *Quaderni di Semantica* 7(1), 49-58.

Fillmore, C. J. (1987). ‘A Private History of the Concept 'Frame'’. In Dirven, R. and Radden, G. (eds.) *Concepts of Case*. Gunter Narr Nverlag Tubingen. 28-36.

Fillmore, C. J. and Atkins, B. T. (1992). ‘Toward a frame-based lexicon: The semantics of RISK and its neighbours’. In Lehrer, A. (ed). *Frames, fields, and contrasts: New essays in semantic and lexical organization*. Routledge. 75-102.

Fillmore, C. J., Petruck, M. R. L., Ruppenhofer, J.and Wright, A. (2003). ‘FrameNet in action: The case of attaching’. *International Journal of Lexicography*, 16(3), 297-332.

Fillmore, C.J. and Baker, C.F. (2001). Frame semantics for text understanding. In *Proceedings of WordNet and Other Lexical Resources Workshop, NAACL*.

Firth J. R. (1957). “A synopsis of linguistic theory 1930-1955”. *Studies in Linguistic Analysis* (special volume of the Philological Society), Oxford, 1-32.

Flowerdew, L. (2003). ‘A combined corpus and systemic-functional analysis of the Problem-Solution pattern in a student and professional corpus of technical writing’. *TESOL Quarterly*, 37(3), 489-511.

Flowerdew, L. (2004). ‘The argument for using English specialized corpora to understand academic and professional settings’. In Connor, U. and Upton, T. (eds.) *Discourse in the Professions: Perspectives from Corpus Linguistics*. Amsterdam: John Benjamins. 11-33.

Foorman, B. R. and Torgesen, J. (2001). ‘Critical elements of classroom and small‐group instruction promote reading success in all children’. *Learning Disabilities Research & Practice*, 16(4), 203-212.

Forsyth, D. R. (1999). *Group Dynamics* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

FrameNet Project website https://framenet.icsi.berkeley.edu/fndrupal/IntroPage.

Fuller, J. M. (2003). 'The influence of speaker roles on discourse marker use'. *Journal of Pragmatics*, 35, 23-45.

Fung, L. and Carter, R. (2007). 'Discourse markers and spoken English: Native and learner use in pedagogic settings'. *Applied Linguistics*, 28(3), 410-439.

Gale, W. A., Church, K. W., and Yarowsky, D. (1992). 'One sense per discourse'. In *Proceedings of the 4th DARPA Speech and Natural Language Workshop*. 233-237.

Gamson, W. A. and Modigliani, A. (1987). 'The changing culture of affirmative action'. In R. G. Braungart & M. M. Braungart (Eds.), *Research in political sociology* (Vol. 3). Greenwich, CT: jAI Press. 137-177.

Gärdenfors, P. (2000). *Conceptual Spaces: the Geometry of Thought*. London: MIT Press.

Gärdenfors, P. (2014). *The Geometry of Meaning: Semantics Based on Conceptual Spaces*. MIT Press.

Gardner, R. (2007). 'The *Right* connections: Acknowledging epistemic progression in talk'. *Language in Society*, 36, 319-341.

Garfinkel, H. (1967). *Studies in Ethnomethodology.* New York: Prentice- Hall.

Gash, H. (2015). 'Knowledge Construction: A Paradigm Shift'. *New Directions for Teaching and Learning*, 2015(143), 5-23.

Gass, S. (1997). *Input, Interaction, and the Second Language Learner*. Mahwah NJ: Lawrence Erlbaum Associate.

Gass, S. and Varounis, E. (1994). 'Input, Interaction and Second Language Production'. *Studies in Second Language Acquisition*. 16, 283-302.

Gass, S., Mackey, A. and Ross‐Feldman, L. (2005). 'Task‐based interactions in classroom and laboratory settings'. *Language Learning*, 55(4), 575-611.

Gatfield, T. (1999). Examining student satisfaction with group projects and peer assessment. *Assessment and Evaluation in Higher Education*, 24(4), 365-378.

Gentner, D and Kurtz, K. J. (2005). 'Relational categories'. In Ahn, W.K., Goldstone, R.L., Love, B.C., Markman, A.B., and Wolff, P.W. (eds.), *Categorization Inside and Outside the Lab*. Washington, DC: American Psychological Association. 151-175.

Gentner, D. (1982). 'Why nouns are learned before verbs: Linguistic relativity versus natural partitioning'. In S. Kuczaj (ed.), Language development: Vol. 2. *Language, Thought and Culture*. Hillsdale, NJ: Erlbaum.

Gentner, D. and Boroditsky, L. (2001). 'Individuation, relativity and early word learning'. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development*. New York: Cambridge University Press.

Gentner, D. and Rattermann, M. J. (1991).' Language and the career of similarity'. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on Language and Thought: Interrelations in Development*. London: Cambridge University Press.

Gergen, J. (1985). 'The social constructionist movement in modern psychology'. *American Psychologist*, 40(3), 266-275.

Gibson, W., Hall, A. and Callery, P. (2006). "Topicality and the structure of interactive talk in face-to-face seminar discussions: implications for research in distributed learning media". *British Educational Research Journal*, 32(1), 77-94.

Giddens, A. (1979). *Central problems in social theory*. Berkeley: University of California Press.

Gillies, R. M. (2003). Structuring cooperative group work in classrooms. *International Journal of Educational Research,* 39(1), 35-49.

Gitlin, T (1980). *The Whole World is Watching: Mass Media in the Making and Unmaking of the New Left*. Berkeley, CA: University of California Press.

Goddard, A. (1998), 'Facing up to market forces'. *Times Higher Education Supplement,* 13, 6-7.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Harmondsworth: Penguin Books.

Goffman, E. (1974). *Frame Analysis: An Essay on the Organization of Experience*. New York: Harper Colophon Books.

Good,T. L. and Grouws, D. A. (1975). 'Teacher rapport: Some stability data'. *Journal of Educational Psychology*, 67(2), 179–82.

Goodson, I.F. and Cole, A.L. (1994). 'Exploring the teacher's professional knowledge: Constructing identity and community'. *Teacher Education Quarterly*, 85-105.

Goodwin, C. (1986). 'Between and within: Alternative sequential treatments of continuers and assessments'. *Human Studies*, 9(2), 205-217.

Goodwin, C. (2007). 'Interactive footing'. In Holt, E. and Clift, R. (eds.) *Reporting Talk: Reported Speech in Interaction*. Cambridge: Cambridge University Press. 16-46.

Goodwin, C. and Goodwin, M. H. (1992). 'Context, activity and participation'. *The Contextualization of Language*. 77-99.

Granger, S. (1993). 'International Corpus of Learner English'. In J. Aarts, P. de Haan and N. Oostdijk (eds.) *English Language Corpora: Design, Analysis and Exploitation*. Amsterdam: Rodopi. 57-71.

Green, J., Franquiz, M. and Dixon, C. (1997), 'The myth of the objective transcript: Transcribing as a situated act'. *TESOL Quarterly,* 31, 172-176.

Gries, S. Th. (2013). '50-something years of work on collocations: What is or should be next…'. *International Journal of Corpus Linguistics*, *18*(1), 137 – 166.

Groom, N. (2007). Phraseology and epistemology in humanities writing: A corpus-driven study. Unpublished PhD dissertation, University of Birmingham, UK.

Gumperz, J. J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.

Gumperz, J. J. (2003). "Interactional Linguistics: A Personal Perspective." In Schiffrin, D., Tannen, D. and Heidi, E. H. (eds). *The Handbook of Discourse Analysis*. Blackwell Publishing.

Gumperz, J. J. and Berenz, N. B. (1993), 'Transcribing conversational exchanges', in J. A. Edwards and M. D. Lampert (eds) *Talking Data: Transcription and Coding in Discourse Research*. Hillsdale, NJ: Lawrence Erlbaum. 91-121.

Gumperz, J.J. (1995). Mutual inferencing in conversation. In Markov á, I., Graumann, C. F. and Foppa, K. (eds.) *Mutualities in Dialogue*. Cambridge University Press. 101-123.

Gumport, P. J. (2000). 'Academic Restructuring: Organizational Change and Institutional Imperatives'. *Higher Education*, 39(1), 67–91.

Gundel, J. K. (1985). 'Shared knowledge and topicality'. *Journal of Pragmatics*, 9, 83-107.

G ünthner, S. (1996). The prosodic contextualization of moral work: an analysis of reproaches in 'why'-formats. *Studies in Interactional Sociolinguistics*, 12, 271-302.

Gupta, M. L. (2004). Enhancing student performance through cooperative learning in physical sciences. *Assessment and Evaluation of Higher Education*, 29(1), 63-73.

Guzmán, A. R. (2004). *Estimation and Evaluation of Plants as Indicators of Tropical Soil Quality from the Knowledge of the Peasants*, Cauca Colombia. Cuvillier Verlag.

Hakulinen, A. and Selting, M. (eds.) *Syntax and Lexis in Conversation*. Amsterdam: Benjamins.

Halliday M.A.K. (1966). "Lexis as a Linguistic level". In C.E. Bazell, J.C. Catford, M.A.K. Halliday and R.H. Robins (eds) *In Memory of J.R. Firth*. Longmans: London. 148-162.

Halliday, M. A. K. (1978). *Language as Social Semiotic. The Social Interpretation of Language and Meaning*. London: Edward Arnold.

Halliday, M. A. K. (1993). 'Some grammatical problems in scientific English'. In M. A. K. Halliday, and J. R. Martin (Eds.), *Writing Science: Literacy and Discursive Power*. London: Falmer Press. 69-85.

Halverson, C. B. and Tirmizi, S. A. (2008). *Effective multicultural teams: Theory and practice. Springer Science*, Business Media B.V.

Hamawand, Z. (2016). *Semantics: A Cognitive Account of Linguistic Meaning*. Equinox Publishing, United Kingdom.

Hancock, M. (1997). Behind classroom code-switching: layering and language choice in L2 learner interaction. *TESOL Quarterly*, 2, 217-235.

Hanks, W. (1996). *Language and Communicative Practice*. Colorado: West View Press.

Hardie, A. (2014). 'Log Ratio: An informal introduction'. Online blog of ESRC Centre for Corpus Approaches to Social Science (CASS) website at Lancaster University, UK. http://cass.lancs.ac.uk/?p=1133.

Hardin, C. D. and Higgins, E. T. (1996). 'Shared reality: How social verification makes the subjective objective'. In Sorrentino, R. M. and Higgins, E. T. (eds.) *Handbook of Motivation and Cognition*, Vol. 3: The interpersonal context. New York, NY, US: Guilford Press. 28-84.

Hartnell-Young, E. A. (2003). ‘Towards knowledge building: Reflecting on teachers’ roles and professional learning in communities of practice’. Unpublished PhD thesis. University of Melbourne.

Hassan, S., Csomai, A., Banea, C., Sinha, R. and Mihalcea, R. (2007). ‘UNT: SubFinder: Combining knowledge sources for automatic lexical substitution’. In *Proceedings of the 4th International Workshop on Semantic Evaluations* (pp. 410-413). Association for Computational Linguistics.

Hata, K. (2016). ‘Contrast-terminal: The sequential placement of trailoff but in extensive courses of action’. *Journal of Pragmatics*, 101, 138-154.

Hayano, K. (2011). Claiming epistemic primacy: *yo*-marked assessments in Japanese. In T. Stivers, L. Mondada, & J. Steensig (Eds.), *The Morality of Knowledge in Conversation*. Cambridge, England: Cambridge University Press. 58-81.

Heinemann, T. (2006). ‘“Will you or can’t you?”: Displaying entitlement in interrogative requests’. *Journal of Pragmatics*, 38, 1081–1104.

Heller, V. (2015). “Academic discourse practices in action: Invoking discursive norms in mathematics and language lessons”. *Linguistics and Education*, 31, 187-206.

Heller, V. and Morek, M. (2015). “Academic discourse as situated practice: An introduction”. *Linguistics and Education*, 31, 174-186.

Henry, A. and Roseberry, R. L. (2001). ‘Using a small corpus to obtain data for teaching a genre’. In Ghadessy, M., Henry, A. and Roseberry, R. L. (eds.) *Small Corpus Studies and ELT: Theory and Practice*, John Benjamins, Amsterdam. 93 – 133.

Heritage, J. (1984). *Garfinkle and Ethomothodology*. Cambridge: Polity Press.

Heritage, J. (1997). ‘Conversation analysis and institutional talk: analyzing data’. In Silverman, D. (ed.), *Qualitative Research: Theory, Method and Practice*. Sage, London, pp. 161–182.

Heritage, J. (2011). Territories of knowledge, territories of experience: Empathic moments in interaction. In T. Stivers, L. Mondada, & J. Steensig (Eds.), *The morality of knowledge in conversation*. Cambridge, England: Cambridge University Press. 159-183.

Heritage, J. (2012). 'Epistemics in action: Action formation and territories of knowledge'. *Research on Language and Social Interaction*, 45(1), 1-29.

Heritage, J. (2013). 'Action formation and its epistemic (and other) backgrounds'. *Discourse Studies*, 15(5), 551-578.

Heritage, J. and Clayman, S. (2010). *Talk in Action. Interactions Identities and Institutions*. Oxford: Wiley-Blackwell.

Heritage, J. and Raymond, G. (2005). 'The terms of agreement: Indexing epistemic authority and subordination in assessment sequences'. *Social Psychology Quarterly*, 68(1), 15-38.

Heritage, J. and Raymond, G. (2012). 'Navigating epistemic landscapes: Acquiescence, agency and resistance in response to polar questions'. In De Ruiter J-P (ed.) *Questions: Formal, Functional and Interactional Perspectives*. Cambridge: Cambridge University Press. 179-192.

Hester, S. and Francis, D. (2000). 'Ethnomethodology, conversational analysis, and 'institutional talk'. *Text*, 20 (3), 391-413.

Hillocks, G. (1986). *Research on written composition: New directions in teaching*. Urbana, IL: ERIC and the National Conference on Research in English.

Hmelo-Silver, C. E. (2003). 'Analyzing collaborative knowledge construction: Multiple methods for integrated understanding'. *Computers & Education*, 41(4), 397-420.

Hmelo-Silver, C. E. (2003). 'Analyzing collaborative knowledge construction: Multiple methods for integrated understanding'. *Computers & Education*, 41(4), 397-420.

Hmelo-Silver, C. E. and Barrows, H. S. (2008). 'Facilitating collaborative knowledge building'. *Cognition and Instruction*, 26(1), 48-94.

Hoey, M. (2005). *Lexical Priming: A new Theory of Words and Language*. London: Routledge.

Houtkoop, H. and Mazeland, H. (1985). 'Turns and discourse units in everyday conversation'. *Journal of Pragmatics*, 9, 595–619.

Howard, R.M. (2001). 'Collaborative pedagogy'. In Tate, G., Taggart, A. R., Schick, K. and Hessler, H.B. (eds.) *A Guide to Composition Pedagogies*. New York: Oxford University Press. 54-70.

Hunston, S. (1995). 'A corpus study of some English verbs of attribution'. *Functions of Language*, 2, 133–158.

Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Hutchby, I. and Wooffitt, R. (2008). *Conversation Analysis*. Cambridge: Polity.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge: MIT Press.

Hyland, K. (1999). 'Academic attribution: Citation and the construction of disciplinary knowledge'. *Applied Linguistics*, 20(3), 341-367.

Hyland, K. (2002). 'Specificity revisited: How far should we go now?'. *English for Specific Purposes*, 21(4), 385-395.

Hyland, K. (2011). 'Academic discourse'. In Hyland, K. and Paltridge, B. (eds.) *Continuum Companion to Discourse Analysis*. London: Continuum. pp 171-184.

Hymes, D. H. (1961). 'The ethnography of speaking'. In Gladwin, T. and Sturtevant, W. C. (eds) *Anthropology and Human Behavior*. Washington, D. C.: Anthropological Society of Washington, D.C. (Reprinted in Fishman, J. A. (ed) (1968). *Readings in the Sociology of Language*. The Hague: Mouton. 99–119.

Imming, P., Sinning, C. and Meyer, A. (2006). Drugs, their targets and the nature and number of drug targets. *Nature reviews Drug discovery*, 5(10), 821-834.

Jaspers, J. (2013). ‘Interactional Linguistics and discourse analysis’. In Gee, J. P. and Handford, M. (eds.) *The Routledge Handbook of Discourse Analysis*. Routledge. 135-146.

Jefferson, G. (1983). ‘On a failed hypothesis: ‘Conjunctionals’ as overlap vulnerable’.Tilburg Papers Lang. Lit. 28, 29--33.

Jefferson, G. (1984). Notes on a Systematic Deployment of the Acknowledgement Tokens 'Yeah' and 'Mm hm'. *Papers in Linguistics*, 17, 197-216.

Jefferson, G. (1985). ‘An exercise in the transcription and analysis of laughter’. In: T.A. van Dijk (ed.), *Handbook of Discourse Analysis (volume III)*. London: Academic Press.

Jefferson, G. (1996). ‘A case of transcriptional stereotyping’. *Journal of Pragmatics*, 26, 159-170.

Jefferson, G. (2004). ‘Glossary of transcript symbols with an introduction’. In Lerner, G. (ed.), *Conversation Analysis: Studies from the First Generation*. Amsterdam, Netherlands: John Benjamins Publishing. 13–31.

Jenks, C.J. (2011). *Transcribing Talk and Interaction: Issues in the Representation of Communication Data*. John Benjamins Publishing.

Jeong, H. and Chi, M. T. (1997). ‘Construction of shared knowledge during collaborative learning’. In Proceedings of the 2nd international conference on Computer support for collaborative learning (130-134). International Society of the Learning Sciences.

Johnson, D. W. and Johnson, R. T. (1999). ‘Making cooperative learning work’. *Theory into Practice*, 38(2), 67-74.

Johnson, D. W., Johnson, R. T. and Smith, K. A. (1998). Cooperative learning returns to college: what evidence is there that it works? *Change*, 30(4), 27-35.

Johnson-Laird, P.N. (1983). *Mental Models: Towards a Cognitive Science of Language,*

Inference, and Consciousness (No. 6). Cambridge, Massachusetts: Harvard University Press.

Jones, S. S. and Smith, L. B. (1993). 'The place of perception in children's concepts'. *Cognitive Development*, 8, 113-139.

Jönsson, L. and Linell, P. (1991). Story generations: From dialogical interviews to written reports in police interrogations. *Text-Interdisciplinary Journal for the Study of Discourse*, 11(3), 419-440.

Jungwirth, H. (1993). "Routines in classroom discourse: an ethnomethodological approach". *Journal of Psychology of Education*, 8(4), 375–387.

Kaartinen, S.and Kumpulainen, K. (2002). 'Negotiating meaning in science classroom communities: Cases across age levels'. *Journal of Classroom Interaction*, 36(2)–37(1), 4–16.

Kamio, A. (1997). *Territory of Information*. Amsterdam: John Benjamins.

Kärkkäinen, E. (2003). *Epistemic stance in English conversation: a description of its interactional functions with a focus on I think*. Philadelphia: John Benjamins Publishing Company.

Kasper, G. (2009). 'Categories, context and comparison in conversation analysis'. In Nguyen, H. and Kasper, G. (eds.) *Talk-in-Interaction: Multilingual Perspectives*. University of Hawaii, National Foreign Language Resource Center, Honolulu. 1--28.

Keil, F. C. (1989). *Concepts, Kinds, and Cognitive Development*. Cambridge, MA: MIT Press.

Kelly, G. and Crawford, T. (1997). 'An ethnographic investigation of the discourse processes of school science'. *Science Education*, 81, 533–559.

Kelly, G. and Green, J. (1998). 'The social nature of knowing: Toward a sociocultural perspective on conceptual change and knowledge construction'. In B. Guzzetti and C. Hynd (Eds.), *Perspectives on Conceptual Change*. Mahwah, NJ: Erlbaum. 145-182.

Kendon, A. (1992). ‘The negotiation of context in face-to-face interaction’. In Duranti, A. and Goodwin, C. (eds.) *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press. 323-334.

Kern, F. and Selting, M. (2013). ‘Conversation Analysis and Interactional Linguistics’. *The Encyclopedia of Applied Linguistics.* 1-5.

Kersten, A.W. and Earles, J. L. (2004). ‘Semantic context influences memory for verbs more than memory for nouns’. *Memory & Cognition*, 32(2), 198-211.

Kidwell, J.J., Vander Linde, K.M. and Johnson, S. L. (2000). ‘Applying corporate knowledge management practices in higher education’. *Educause Quarterly*, 4, 28-33.

Kilgarriff, A. (2005). ‘Language is never, ever, ever, random’, *Corpus Linguistics and Linguistic Theory,* 1(2), 263–276.

Kimmel, K. and Volet, S. (2010). “Significance of context in university students’ (meta)cognitions related to group work: A multi-layered, multi-dimensional and cultural approach”. *Learning and Instruction*, 20(6), 449-464.

Kittay, E. F. and Lehrer, R. (eds.) (2012). *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization*. London: Routledge.

Kittleson, J. M. and Southerland, S. A. (2004). ‘The role of discourse in group knowledge construction: A case study of engineering students’. *Journal of Research in Science Teaching*, 41(3), 267-293.

Knight, D. (2009). ‘A multi-modal corpus approach to the analysis of backchanneling behaviour’ Unpublished PhD dissertation, University of Nottingham, UK.

Kolb, D. A. (1981). ‘Learning styles and disciplinary differences’. *The Modern American College*, 232-255.

Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J. and Ryan, M. (2003). 'Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design (tm) into practice'. *The Journal of the Learning sciences*, 12(4), 495-547.

Komatsu, L.K. (1992). 'Recent views of conceptual structure'. *Psychological Bulletin*, 112(3), 500-526.

Koschmann, T., Glenn, P. and Conlee, M. (2000). 'When is a problem-based tutorial not tutorial? Analyzing the tutor's role in the emergence of a learning issue'. In D. Evensen and C. E. Hmelo (eds.), *Problem-based Learning: A Research Perspective on Learning Interactions*. Mahwah NJ: Erlbaum. 53-74.

Koshik, I. (2005). *Beyond Rhetorical Questions: Assertive Questions in Everyday Interaction*. Amsterdam: John Benjamins.

Koteyko, N. (2006). 'Corpus linguistics and the study of meaning in discourse'. *The Linguistics Journal*, 1(2), 131-154.

Krajcik, J., McNeill, K. L. and Reiser, B. J. (2008). 'Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy'. *Science Education*, 92(1), 1-32.

Kramsch, C. (2002). 'Introduction: "How can we tell the dancer from the dance?"' In Kramsch, C. (ed.) *Language Acquisition and Language Socialization*. London: Continuum. 1-30.

Kretzschmar, W. A., Meyer, C. F. and Ingegneri, D. (1997). 'Uses of inferential statistics in corpus studies'. In M. Ljung (Ed.), *Corpus-based Studies in English: Papers from the Seventeenth International Conference on English Language Research on Computerized Corpora* (ICAME 17), Stockholm, May 15–19, 1996. Amsterdam: Rodopi. 167-177.

Laal, M. (2011). 'Knowledge management in higher education'. *Procedia Computer Science*, 3, 544-549.

Labov, W. and Fanshel, D. (1997). *Therapeutic Discourse: Psychotherapy as Conversation*. New York: Academic Press.

Lakoff, G. (1977). Linguistic gestalts. In Papers from the Regional Meeting. Chicago Ling. Soc. Chicago, Ill. (Vol. 13). 236-287.

Lakoff, G. (1987a). 'Cognitive models and prototype theory'. In U. Neisser (Ed.), *Concepts and Conceptual Development*. New York: Cambridge University Press. 63-100.

Lakoff, G. (1987b). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: Univ. Chicago Press.

Lakoff, G. (1993). 'The contemporary theory of metaphor'. In Ortony, A. (ed.) *Metaphor and Thought* (2nd edition). Cambridge: Cambridge University Press. 205-251.

Lakoff, G. and Johnson, M. (1980a). 'The metaphorical structure of the human conceptual system'. *Cognitive Science*, 4(2), 195-208.

Lakoff, G. and Johnson, M. (1980b). *Metaphor We Live By*. Chicago: University of Chicago Press.

Lakoff, R. (1989). 'The limits of politeness'. *Multilingua*, 8(2-3), 101-129.

Lakoff, G. and Turner, M. (1989). *More Than Cool Reaso*n. Chicago: University of Chicago Press.

Langacker, R.W. (1987). *Foundations of Cognitive Grammar: Theoretical Prerequisites* (Vol. 1). Stanford, California: Stanford University Press.

Lantolf, J. P. and Thorne, S. L. (2000). 'Sociocultural theory and second language learning'. In Lantolf, J. P. (ed.) *Sociocultural Theory and Second Language Learning*. Oxford University Press. 197-221.

Lave, J. (1988). *Cognition in Practice*. Cambridge: Cambridge University Press.

Lave, J. and Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge, UK: Cambridge University Press.

Lee, D. (2002). *Cognitive linguistics: An introduction*. Oxford, UK: Oxford University Press.

Leech, G. (1997). 'Introducing Corpus Annotation'. In Garside, R., Leech, G., and McEnery, A. (eds.), C*orpus Annotation: Linguistic Information from Computer Text Corpora*. Longman, London. 1–18.

Leech, G. and Fallon, R. (1992). 'Computer corpora — what do they tell us about culture?', *ICAME Journal*, 16, 29–50.

Leech, G. and Smith, N. (1999). 'The use of tagging'. In van Halteren, H, (ed.) *Syntactic Wordclass Tagging*, Kluwer, The Netherlands. 23 – 36.

Lencioni, P. (2006). *The Five Dysfunctions of a Team*. San Francisco: John Wiley & Sons.

Lerner, G. H. and Kitzinger, C. (2007). 'Extraction and aggregation in the repair of individual and collective self-reference'. *Discourse Studies*, 9, 526-557.

Levinson, S. (2012). Interrogative intimations: On a possible social economics of interrogatives. In: J.-P. De Ruiter (ed.), *Questions: Formal, functional and interactional perspectives*. Cambridge: Cambridge University Press, 11-32.

Liddicoat, A. J. (2007). *An Introduction to Conversation Analysis*. London: Continuum.

Liebscher G. and Dailey-O'Cain, J. (2003) Conversational repair as a role-defining mechanism in classroom interaction. *The Modern Language Journal*, 87, 375-390.

Linell, P. (1998a). 'Contexts in discourse and discourse in context'. In Linell, P. *Approaching Dialogue: Talk, Interaction and Context in Dialogical Perspectives*. Amsterdam: J. Benjams. Pub. Co. 127-158.

Linell, P. (1998b). ‘Discourse across boundaries: On recontextualizations and the blending of voices in professional discourse’. *Text-Interdisciplinary Journal for the Study of Discourse*, 18(2), 143-158.

Linell, P. (2008). ‘Perspectives on language and discourse’. In Linell, P. *Approaching Dialogue: Talk, Interaction and Context in Dialogical Perspectives*. Amsterdam: J. Benjamins Pub. Co. 3-15.

Linell, P. and Luckmann, T. (1991). ‘Asymmetries in dialogue: Some conceptual preliminaries. In Markova, I. and Foppa, K. (eds.) *Asymmetries in Dialogue*. Hemel Hempstead: Harvester Wheatsheaf. 1-20.

Linn, M. C. and Burbules, N. C. (1993). ‘Construction of knowledge and group learning’. In K. Tobin and D. Tippins (eds.), *The Practice of Constructivism in Science Education*. Hillsdale, NJ: Erlbaum. 91-120.

Linnenbrink, E. A. and Pintrich, P. R. (2001). ‘Multiple goals, multiple contexts: The dynamic interplay between personal goals and contextual goal stresses’. In Volet, S. and Järvelä, S. (eds.). *Motivation in Learning Contexts: Theoretical Advances and Methodological Implications*. Elmsford, NY, US: Pergamon Press. 251-269.

Litkowski, K., (2007). ‘CLR: integration of FrameNet in a text representation system. In *Proceedings of the 4th International Workshop on Semantic Evaluations* (pp. 113-116). Association for Computational Linguistics.

Long, M. (1996). ‘The Role of the Linguistic Environment in Second Language Acquisition’. In W. Ritchie and T. Bhatia (eds.), *Handbook of Research on Second Language Acquisition*. New York, Academic. 413- 468.

Long, M. and Robinson, P. (1998). 'Theory, Research and Practice'. In C. Doughty & J.Williams (eds.), *Focus on Form in Classroom Second Language Acquisition*. Cambridge, England: Cambridge University Press. 15-41.

Loock, R. and O'Connor, K.M. (2013). 'The discourse functions of nonverbal appositives'. *Journal of English Linguistics*, 41(4), 332-358.

Louw, B. (1993). 'Irony in the text or insincerity in the writer? — the diagnostic potential of semantic prosodies'. In M. Baker, G. Francis and E. Tognini-Bonelli (eds.), *Text and Technology: in Honor of John Sinclair*. Amsterdam: Benjamins. 157-176.

Louw, B. (2000). 'Contextual prosodic theory: Bringing semantic prosodies to life' in C. Heffer, H. Sauntson, and G. Fox (eds). *Words in Context: A Tribute to John Sinclair on his Retiremen*t. Birmingham: University of Birmingham.

Lowe, J. B., Baker, C. F. and Fillmore, C. J. (1997). 'A frame-semantic approach to semantic annotation'. Proceedings of the SIGLEX workshop 'Tagging Text with Lexical Semantics: Why, What and How?' April 1997, Washington DC.

Lucy, J. A. (1997). 'Linguistic relativity'. *Annual Review of Anthropology*, 291-312.

Lukes, S. (1974). *Power: A Radical View*. London: Macmillan.

Macdonald, J. (2003). 'Assessing online collaborative learning: process and product'. *Computers & Education*, 40(4), 377-391.

Mackey, A. (1999). 'Input, Interaction and Second Language Development. An Empirical Study of Question Formation in ESL'. *Studies in Second Language Acquisition*. 21, 557-587.

MacLachlan, G. and Reid, I. (1994). *Framing and Interpretation*. Carlton, Victoria: Melbourne University Press.

Mahlberg, M. (2007). 'Clusters, key clusters and local textual functions in Dickens'. *Corpora*, 2(1), 1–31.

Makitalo, A. and Saljo, R. (2001) 'Talk in institutional context and institutional context in talk: Categories as situated practices'. *Text*, 22(1), 57-82.

Margolis, J. S. (1992). Piranhas, monsters, and jugglers: the psychology of gender and academic discourse. *On Teaching and Learning, 4, 1992.* Harvard University: Book Center on Teaching at Harvard University.

Markee, N. P. (1995). "Teachers' answers to students' questions: problematizing the issue of making meaning". *Issues in Applied Linguistics*, 6(2), 63–92.

Marx, K. and Engels, F. (1845). *German Ideology*. London: Routledge and Kegan Paul.

Mason, L. (1998). 'Sharing Cognition to Construct Scientific Knowledge in School Context: The Role of Oral and Written Discourse', *Instructional Science,* 26(5), 359–89.

Maynard, D. (2003). *Bad news, Good news: Conversational Order in Everyday Talk and Clinical Settings*. Chicago, IL: University of Chicago Press.

Mayordomo, R. M. and Onrubia, J. (2015). 'Work coordination and collaborative knowledge construction in a small group collaborative virtual task'. *The Internet and Higher Education*, 25, 96-104.

McCarthy, D. (2001). 'Lexical Acquisition at the Syntax- Semantics Interface: Diathesis Aternations, Subcategorization Frames and Selectional Preferences'. Unpublished Ph.D. thesis, University of Sussex.

McCarthy, D. (2002). 'Lexical substitution as a task for WSD evaluation'. In Proceedings of the ACL Workshop on Word Sense Disambiguation: Recent Successes and Future Directions, Philadelphia.

McCarthy, M. (2003). Talking back: "Small" interactional response tokens in everyday conversation. *Research on Language and Social Interaction,* 36(1), 33-63.

McCarthy, M. J. and Carter, R. A. (1997). 'Written and spoken vocabulary'. In Schmitt, N. and McCarthy, M. J. (eds.) *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press. 20-39.

McEnery, T. (2006a). *Swearing in English: Bad Language, Purity and Power from 1586 to the Present.* Abington, UK: Routledge.

McEnery, T. (2006b). 'The moral panic about bad language in England, 1691 – 1745'. *Journal of Historical Pragmatics,* 7(1), 89-113.

McEnery, T. (2015). Personal communication in Lancaster Summer School for Corpus Linguistics Approaches to Social Sciences, Lancaster University.

McEnery, T. and Andrew Wilson, A. (2001). *Corpus Linguistics*. Edinburgh : Edinburgh University Press.

Medin, D. L. and Smith, E. E. (1984). 'Concepts and concept formation'. *Annual Review of Psychology*, 35(1), 113-138.

Medin, D.L. (1989). 'Concepts and conceptual structure'. *American Psychologist*, 44(12), 14-69.

Mehan, H. (1993). 'Beneath the skin and between the ears'. In Lave, J. and Chaiklin, S. (eds.) *Understanding Practice: Perspectives on Activity and Context*. Cambridge: Cambridge University Press. 241–268.

Mercer, N. (1995). *The Guided Construction of Knowledge: Talk amongst Teachers and Learners*. Clevedon, Avon, England; Philadelphia: Multilingual matters.

Mihalcea, R. (2006). Knowledge based methods for word sense disambiguation. In E. Agirre and P. Edmonds (Eds.) *Word Sense Disambiguation: Algorithms and applications., volume 33 of Text, Speech and Language Technology.* Springer.

Mihalcea, R. and Moldovan, D. (2001). ‘Extended wordnet: Progress report’. In Proceedings of NAACL Workshop on WordNet and Other Lexical Resources, Pittsburgh, PA.

Minsky, M. (1969). ‘Form and Content in Computer Science’. *Journal of the Association for Computing Machinery*, 17(2), 1-38.

Minsky, M. (1974). ‘A framework for representing knowledge’. *Artificial Intelligence*, 306, 1-82.

Minsky, M. (1975). ‘A framework for representing knowledge’. In Winston, P. H. (ed.) *The Psychology of Computer Vision*. McGraw-Hill, New York. 211–277.

Minsky, M. (1977). ‘Frame-system theory’. In Johnson-Laird, P. and Wason, P. C. (eds.) *Thinking Readings in Cognitive Science*. Cambridge: Cambridge University Press. 355-376.

Mishler, E. G. (1991), 'Representing discourse: The rhetoric of transcription'. J*ournal of Narrative and Life History*, 1, 255-80.

Moje, E.B. and Shepardson, D. P. (1998). ‘Social interactions and children’s changing understanding of electric circuits: Exploring unequal power relations in ‘‘peer’’-learning groups’. In Guzzetti, B. and Hynd. C. (eds.), *Perspectives on Conceptual Change*. Mahwah, NJ: Erlbaum. 225-234.

Moon, R. (2010). ‘What can a corpus tell us about lexis’. In O’Keeffe, A. and McCarthy, M. (eds.) *The Routledge Handbook of Corpus Linguistics*. London: Routledge. 197-211.

Moore, W. S. (2002). ‘Understanding learning in a postmodern world: Reconsidering the Perry scheme of intellectual and ethical development’. In Hofer, B. K. and Pintrich, P. R. (eds). *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing.* Mahwah, NJ: Lawrence Erlbaum Associates. 17-36.

Näykki, P. and Järvelä, S. (2008). 'How pictorial knowledge representations mediate collaborative knowledge construction in groups'. *Journal of Research on Technology in Educatio*n, 40(3), 359-387.

Needham, R. (1987). 'Teaching strategies for developing understanding in science'. Children's Learning in Science Project, Centre for Studies in Science and Mathematics Education, University of Leeds, UK.

Nerlich, B. and Clarke, D. D. (2000). "Semantic fields and frames: Historical explorations of the interface between language, action and cognition". *Journal of Pragmatics*, 32, 125-150.

Neumann, R., Parry, S. and Becher, T. (2002). 'Teaching and learning in their disciplinary contexts: A conceptual analysis'. *Studies in Higher Education*, 27(4), 405-417.

Newell, A. and Simon, H. A. (1972). *Human Problem Solving* (Vol. 104, No. 9). Englewood Cliffs, NJ: Prentice-Hall.

Newton, I. (2009). 'Rules for the study of natural philosophy'. *The general scholium. Philosophy of Science: An Historical Anthology*, 794-796.

Nikander, P. (2008). 'Working with Transcripts and Translated Data'. *Qualitative Research in Psychology*, 5(3), 225-231.

Nishina, Y. (2010). 'Evaluative meanings and disciplinary values: A corpus-based study of adjective patterns in research articles in Applied Linguistics and Business Studies'. Unpublished PhD dissertation, University of Birmingham, UK.

Norman, D. A. (1972). Memory, knowledge, and the answering of questions. Paper presented at *the Loyola Symposium on Cognitive Psychology*. Chicago.

Norman, D. A. (1988). *The Psychology of Everyday Things*. New York: Basic Books.

North, S. (2005). 'Disciplinary variation in the use of Theme in undergraduate essays'. *Applied Linguistics*, 26(3), 431-452.

Nunan, D. (1988). *The Learner Centred Curriculum*. Cambridge: Cambridge University Press.

Nunan, D. (1992). *Research Methods in Language Learning*. Cambridge: Cambridge University Press.

O'Connor, K.E. (2008). "You choose to care": Teachers, emotions and professional identity. *Teaching and Teacher Education*, 24(1), 117-126.

O'Keeffe, A., and Walsh, S. (2012). Applying corpus linguistics and conversation analysis in the investigation of small group teaching in higher education. *Corpus Linguistics and Linguistic Theory*, 8(1), 159–181.

O'Keeffe, A., McCarthy, M. J. and Carter, R. A. (2007). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge University Press.

Olsen, J. P. (2007). 'The institutional dynamics of the European university'. In Maassen, P. A. and Olsen, J. P. (eds.) *University Dynamics and European Integration* (Vol. 4). Dordrecht: Springer. 25-54.

Ono, T. and Couper-Kuhlen, E. (2007). 'Increments in cross-linguistic perspective: Introductory remarks'. *Pragmatics*, 17, 505-512.

Paavola, S., Lipponen, L. and Hakkarainen, K. (2004). 'Models of innovative knowledge communities and three metaphors of learning'. *Review of Educational Research*, 74(4), 557-576.

Palincsar, A. S. (1998). 'Social constructivist perspectives on teaching and learning'. *Annual Review of Psychology*, 45, 345–375.

Pan, Y. (2013). 'Interactional Linguistics as a research perspective'. In Carol, A. C. (ed) *The Encyclopedia of Applied Linguistics*. Wiley Online Library (Online service).

Pan, Z. and Kosicki, G. M. (1993). 'Framing analysis: An approach to news discourse'. *Political Communication*, 10, 55-75.

Park, I. (2012). Asking different types of polar questions: The interplay between turn, sequence, and context in writing conferences. *Discourse Studies*, 14 (5), 613-633.

Partington, A. (1998). *Patterns and Meaning. Using Corpora for English Language Research and Teaching*. Amsterdam: Benjamins.

Partington, A. (2004). "Utterly content in each other's company": semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, 9(1), 131–156.

Pecina, P. (2010). 'Lexical association measures and collocation extraction'. *Language Resources and Evaluation*, *44*(1-2), 137 – 158.

Peirce, C. S. (1934). 'Pragmatism and pragmaticism'. In C. Hartshorne and P. Weiss (eds.), *The Collected Papers of Charles Sanders Peirce (1931-1935) Volume 5*. Cambridge, Mass: Harvard University Press.

Peschl, M. F. (2010). 'Knowledge construction and knowledge creation as key factors in educational and innovation processes'. In Risku, H. and Peschl, M. F. (eds.) *Kognition und Technologie im kooperativen Lernen. Vom Wissenstransfer zur Knowledge Creation* .15-31.

Petruck, M.R. (1996). 'Frame semantics and the lexicon: Nouns and verbs in the body frame'. In Shibatani, M. and Thompson, S. (eds.) *Essays in semantics and pragmatics: In honor of Charles J. Fillmore*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 279-297.

Petruck, M.R., Fillmore, C.J., Baker, C.F., Ellsworth, M. and Ruppenhofer, J. (2004). Reframing framenet data. In Proceedings of *The 11th EURALEX International Congress*. 405-416.

Phillips, M. (1983). *Lexical macrostructure in science text* (Unpublished doctoral dissertation). University of Birmingham, Birmingham, UK.

Phillips, M. (1985). *Aspects of Text Structure: An Investigation of the Lexical Organisation of Text*. Amsterdam, Netherlands: North-Holland.

Phillips, M. (1989). 'Lexical structure of text'. *English Language Research*, 12.

Phillips, M. (1989). *Lexical Structure of Text* [Discourse Analysis Monograph 12]. University of Birmingham, Birmingham, UK.

Piaget, J. (1972). *Psychology and Epistemology: Towards a Theory of Knowledge*. Translated from French by PA Wells. A. Lane. Chicago: Penguin.

Pica, T. (1994). 'Research on Negotiation: What Does it Reveal about Second Language Learning Conditions, Processes and Outcomes?'. *Language Learning*, 44(3), 493-527.

Piirainen-Marsh, A. and Tainio, L. (2014). 'Asymmetries of knowledge and epistemic change in social gaming interaction'. *Modern Language Journal*, 98 (4), 1022-1038.

Popov, V., Brinkman, D., Biemans, H. J. A., Mulder, M., Kuznetsov, A. and Noroozi, O. (2012). "Multicultural student group work in higher education: An explorative case study on challenges as perceived by students". *International Journal of Intercultural Relations*. 36(2), 302-317.

Port, R.F. and Van Gelder, T. (eds.) (1995). *Mind as Motion: Explorations in the Dynamics of Cognition*. London: MIT press.

Potter, J. (1996). *Representing Reality: Discourse, Rhetoric and Social Construction*. London: Sage.

Potter, J. and Edwards, D. (2001).'Discursive social psychology'. In Robinson, W.P. and Giles, H. (eds.) *The New Handbook of Language and Social Psychology*, *9*, Chichester, England: J. Wiley. 103-118.

Prentice, S. (2010). 'Using automated semantic tagging in Critical Discourse Analysis: A case study on Scottish independence from a Scottish nationalist perspective'. *Discourse & Society*, 21(4), 405-437.

Psathas, G. and Anderson, T. (1990). The'practices' of transcription in conversation analysis. *Semiotica*, 78(1-2), 75-100.

Putnam, H. (1975a). 'Is semantics possible?'. In H. Putnam (ed.), *Philosophical Papers: Volume 2. Mind, Language and Reality*. Cambridge, England: Cambridge University Press. 139-152.

Putnam, H. (1975b). The meaning of "meaning." In H. Putnam (Ed.*), Philosophical Papers: Volume 2. Mind, Language and Reality*. Cambridge, England: Cambridge University Press. 215-271.

Putnam, L. L., and Holmer, M. (1992). 'Framing, reframing, and issue development'. In Putnam, L. L. and Roloff, M. E. (eds.) *Communication and Negotiation*. Newbury Park, CA: Sage. 128-155.

Quantz, R. A. (2001). "On seminars, rituals and cowboys". *Teacher's College Record*, 103(5), 896–922.

Rampton, B. (1999), '*Deutsch* in Inner London and the animation of an instructed foreign language'. *Journal of Sociolinguistics,* 3, 480-504.

Raymond, G. (2003). 'Grammar and social organization: Yes/no interrogatives and the

structure of responding. *American Sociological Review*, 68, 939-967.

Raymond, G. and Heritage, J. (2006). 'The epistemics of social relations: Owning grandchildren'. *Language in Society*, 35, 677–705.

Rayson, P. (2002). 'Matrix: A statistical method and software tool for linguistic analysis through corpus comparison'. Unpublished PhD dissertation, Lancaster University, UK.

Rayson, P. (2008). 'From key words to key semantic domains', *International Journal of Corpus Linguistics,* 13(4), 519-549.

Rayson, P. (2016). Personal communication on the use of Wmatrix corpus analysis software.

Rayson, P. and Wilson, A. (1996). The ACAMRIT semantic tagging system: progress report. In Evett, L. J., and Rose, T. G. (eds.) Language Engineering for Document Analysis and Recognition (LEDAR), AISB96 Workshop proceedings, Brighton, England. Faculty of Engineering and Computing, Nottingham Trent University, UK. 13 – 20.

Rayson, P., Archer, D., Piao, S. and McEnery, T. (2004) 'The UCREL Semantic Analysis System', in Proceedings of the Workshop on Beyond Named Entity Recognition Semantic Labelling for NLP Tasks in Association with the 4th International Conference on Language Resources and Evaluation (LREC), pp. 7–12. Available at: http://eprints.lancs.ac.uk/12453/1/usas_lrec04ws.pdf

Reddy, M. (1993). 'The conduit metaphor'. In Ortony, A. (ed.) *Metaphor and thought*. Cambridge: Cambridge University Press. 284-324.

Resnick, M. (1987). 'The 1987 Presidential Address learning in school and out'. *Educational Researcher*, 16(9), 13-54.

Rey, G. (1983). 'Concepts and stereotypes'. *Cognition*, 15(1), 237-262.

Richmond, G. and Striley, J. (1996). 'Making meaning in classrooms: Social processes in small-group discourse and scientific knowledge building'. *Journal of Research in Science Teaching*, 33(8), 839-858.

Riesman, D. (1998). *On Higher Education: The Academic Enterprise in an Era of Rising Student Consumerism*. Somerset, NJ: Transaction Publishers.

Rommetveit, R. (1985). 'Language acquisition as increasing linguistic structuring of experience and symbolic behaviour control'. In Wertsch, J. V. (ed.) *Culture, Communication, and Cognition: Vygotskian Perspectives*. Cambridge: Cambridge University Press. 183-204.

Rommetveit, R. (1992). 'Outlines of a dialogically based social-cognitive approach to human cognition and communication'. In Wold, A. H. (ed). *The Dialogical Alternative: Towards a Theory of Language and Mind*. Oslo: Scandinavian University Press.19-44.

Rorty, R., Williams, M. and Bromwich, D. (1980). *Philosophy and the Mirror of Nature* (Vol. 401). Princeton, NJ: Princeton University Press.

Roschelle, J. and Teasley, S. D. (2012). 'The construction of shared knowledge in collaborative problem solving'. In O'Malley, C. (ed.) *Computer Supported Collaborative Learning*. Springer Berlin Heidelberg. 69-97.

Rosenfeld, B. and Penrod, S.D. (2011). *Research Methods in Forensic Psychology*. John Wiley and Sons.

Roth, W. M. (1993). 'Construction sites: Science labs and classrooms'. In K.Tobin&D. Tippins (eds.) *The Practice of Constructivism in Science Education*. Hillsdale, NJ: Erlbaum. 145-170.

Rowley, J. (2000). 'Is Higher Education Ready for Knowledge Management?'. The *International Journal of Educational Management*, 14(7). 325-333.

Rumelhart, D.E. (1980). ‘Schemata: The building blocks of cognition’. In. Spiro, R., Bruce, B. and Brewer, W. (eds.) *Theoretical Issues in Reading Comprehension: Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence, and Education*. Routledge.

Ruppenhofer, J., Ellsworth, M., Petruck, M.R., Johnson, C.R. and Scheffczyk, J. (2006). FrameNet II: Extended theory and practice.

Ryle, G. (1949). *The Concept of Mind*. Harmondsworth, Middlesex: Penguin Books.

Sacks, H. (1984). ‘Notes on methodology’. In: J.M. Atkinson and J. Heritage (eds), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press. 21-27.

Sacks, H. (1992). In Jefferson, G. (ed.) *Lectures on Conversation*, vol.II. Oxford: Blackwell.

Sacks, H., Schegloff, E. A. and G. Jefferson (1974). ‘A simplest systematics for the organization of turn-taking for conversation’. *Language*, 50(4), 696-735.

Salzmann, Z. (1993). *Language, Culture and Society: An Introduction to Linguistic Anthropology*. Boulder, Colorado: Westview Press.

Sandra, K.A.T.Z. (2009). ‘Knowledge co-construction and initiative in peer learning interactions’. In Dimitrova, V., Mizoguchi, R., du Boulay, B. and Graesser, A. (eds.) *Artificial Intelligence in Education: Building Learning Systems that Care: from Knowledge Representation to Affective Modelling*. Amsterdam: IOS Press. 325-332.

Sarangi, S. and Roberts, C. (1999). *Talk, Work and Institutional Order: Discourse in Medical, Mediation and Management Settings (Vol. 1).* Walter de Gruyter.

Sarangi, S. and Slembrouck, S. (1996). *Language, Bureaucracy and Social Control*. London: Longman.

Scardamalia, M. and Bereiter, C. (2006). ‘Knowledge building: Theory, pedagogy, and technology’. In K. Sawyer (ed.) *Cambridge Handbook of the Learning Sciences*. New York:

Cambridge University Press. 97-118.

Schank, R.C. (1972). ‘Conceptual dependency: A theory of natural language understanding’. *Cognitive Psychology*, 3(4), 552-631.

Schank, R.C. and Abelson, R.P. (1975). *Scripts, Plans, and Knowledge*. New Haven, CT: Yale University. 151-157.

Schegloff, E. (1982). ‘Discourse as interactional achievement: Some uses of “uh huh” and other things that come between sentences’. In Tannen, D. (ed.), *Analyzing Discourse, Text, and Talk*. Washington, DC: Georgetown University Press. 71-93.

Schegloff, E. (1991). ‘Reflections on talk and social structure’. In Boden, D. and Zimmerman, D. (eds.) *Talk and Social Structure. Studies in Ethnomethodology and Conversation Analysis.* Cambridge: Polity Press. 44-70.

Schegloff, E. (1992). ‘In another context’. In Duranti, A. and Goodwin, C. (eds.) *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press. 191-227.

Schegloff, E. A. (1972). “Notes on a conversational practice: Formulating place”. *Studies in Social Interaction*, 75, 75-119.

Schegloff, E. A. (1984). ‘On some questions and ambiguities in conversation’. In J. M. Atkinson and J. Heritage (eds.) *Structures of Social Action*. Cambridge, England: Cambridge University Press. 28-52.

Schegloff, E. A. (1996a), 'Confirming allusions: Towards an empirical account of action'. *American Journal of Sociology*, 104, 161-216.

Schegloff, E. A. (1996b), 'Turn organization: One intersection of grammar and interaction'. In E. Ochs, E. A. Schegloff and S. A. Thompson (eds), *Interaction and Grammar.* Cambridge: Cambridge University Press. 52-133.

Schegloff, E.A. (1997). ‘Practices and actions: boundary cases of other-initiated repair’. *Discourse Processes*, 23, 499-545.

Schiappa, E. (2003). *Defining reality: Definitions and the politics of meaning*. Carbondale and Edwardsville: Southern Illinois University Press.

Schiffrin, D. (1987). *Discourse Markers*. Cambridge: Cambridge University Press.

Schleppegrell, M. (2001). ‘Linguistic features of the language of schooling’. *Linguistics and Education*, 12(4), 431–459.

Schourup, L. (1999). ‘Discourse markers: tutorial overview’. *Lingua*, 107, 227-265.

Schraw, G. and Olafson, L. (2003).’Teachers' epistemological world views and educational practices’. *Journal of Cognitive Education and Psychology*, 3(2), 178-235.

Schwartz, S. P. (1978). ‘Putnam on artifacts’. *Philosophical Review*, 87, 566-574.

Schwartz, S. P. (1979). ‘Natural kind terms’. *Cognition*, 7, 301-315.

Scott, M. (1999). *WordSmith Tools Help Manual*. Version 3.0. Oxford, UK: Mike Scott ad Oxford University Press.

Scott, M. (2000b). ‘Reverberations of an Echo’. In Lewandowska-Tomaszczyk, B. and Melia, P. J. (eds.) *PALC’99: Practical applications in language corpora: papers from the International conference at the University of Łódź*, 15-18. Peter Lang, Frankfurt. 49–65.

Scott, M. (2001). ‘Mapping key words to problem and solution’. In Scott, M. and Thompson, G. (eds.) *Patterns of Text: In Honour of Michael Hoey*. Benjamins, Amsterdam. 109 –127.

Scott, M. (2009). ‘In search of a bad reference corpus’. In Archer, D. (ed.) *What’s in a Word List? Investigating Word Frequency and Keyword Extraction*. Oxford: Ashgate. 79-92.

Scott, M. (2010). ‘What can corpus software do?’ In A. O’Keeffe and M. McCarthy (eds.) *The Routledge Handbook of Corpus Linguistics*. Oxfordshire: Routledge. 136-151.

Scott, M. (2015). WordSmith Tools. Version 6.0. Oxford, UK: Mike Scott ad Oxford University Press.

Scott, P. H., Mortimer, E. F. and Aguiar, O. G. (2006). 'The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons'. *Science Education*, 90, 605–631.

Seedhouse, P. (2004a). 'Conversation analysis methodology'. *Language Learning*, *54*(S1), 1-54.

Seedhouse, P. (2004b). *The interactional architecture of the language classroom: a conversation analysis perspective*. Oxford: Blackwell Publishing.

Selting, M. (2008). 'Linguistic resources for the management of interaction'. In G. Antos, G., Ventola, E. and Weber, T. (eds.), *Handbook of Interpersonal Communication*. Berlin: Walter de Gruyter. 217-253.

Semantics of Business Vocabulary and Business Rules (SBVR) (2015). (version 1.3) http://www.omg.org/spec/SBVR/1.3/.

Sennett, R. (1977). *The Fall of Public Man*. New York: Knopf.

Seo, M. S. (2008). 'A conversation-analytic study of repair practices in one-on-one ESL tutoring'. Unpublished PhD dissertation, University of Illinois at Urbaba-Champaign.

Sinclair J. M. (1991). *Corpus, Concordance, Collocatio*n. Oxford: Oxford University Press.

Sinclair, J. (1995). 'Corpus typology – a framework for classification'. In Melchers, G. and Warren, B. (eds.) *Studies in Anglistics*. Almqvist & Wiksell, Stockholm. 17–33.

Sinclair, J. (1996). 'The search for units of meaning'. *Textus* online only, 9(1), 25-49.

Sinclair, J. (2001). 'Preface to small corpus studies and ELT'. In Ghadessy, M., Henry, A. and Roseberry, R. (eds.) *Small Corpus Studies and ELT*. Amsterdam: John Benjamins. vii-xv.

Sinclair, J. M. (2004). *Trust the Text: Language, Corpus and Discourse*. London: Routledge.

Siyanova, A. and Schmitt, N. (2008). 'L2 learner production and processing of collocation: A multi-study perspective'. *Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes, 64*(3), 429 – 458.

Slimani, A. (1989).'The role of topicalization in classroom language learning'. *System,* 17(2), 223-234.

Smadja, F. (1993). 'Retrieving collocations from text: Xtract'. *Computational Linguistics,* 19(1), 143 – 177.

Smagorinsky, P. and Fly, P. K. (1993). 'The social environment of the classroom: A Vygotskian perspective on small group process'. *Communication Education,* 42(2), 159-171.

Smith, S.W. and Jucker, A.H. (1998). 'Interactive aspects of reference assignment in conversations'. *Pragmatics & Cognition, 6*(1), 153-187.

Spuzic, S. and Nouwens, F. (2004). 'A Contribution to defining the term 'Definition''. Iss*ues in Informing Science and Information Technology-InSITE2004*.

Stahl, G. (2006). *Group Cognition: Computer Support for Building Collaborative Knowledge*. Cambridge, MA: MIT Press.

Stichweh, R. (2003). 'Differentiation of scientific disciplines: causes and consequences'. *Encyclopedia of Life Support Systems (EOLSS)*. 1-8.

Stivers, T. (2005). 'Modified repeats: One method for asserting primary rights from second position'. *Research on Language and Social Interaction,* 41(1), 297-321.

Stivers, T. and Rossano, F. (2010). Mobilizing response. *Research on Language and Social Interaction,* 43, 3–31.

Stivers, T., Mondada, L. and Steensig, J. (2011). 'Knowledge, morality and affiliation in social interaction'. In Stivers, T., Mondada, L. and Sttensig, J. (eds.) *The Morality of Knowledge in Conversation. Cambridge, England: Cambridge University Press*. 3-26.

Stokoe, E. H. (2000). 'Constructing topicality in university students' small-group discussion: A conversation analytic approach'. *Language and Education*, 14(3), 184-203.

Strawson, P. E. (1964). 'Identifying reference and truth-values'. *Theoria*, 30, 96-118.

Stubbs, M. (1995). 'Collocations and semantic profiles'. *Functions of Language*, 2(1), 23-55.

Stubbs, M. (2010). 'Three concepts of keywords' in Bondi, M. and Scott, M. (eds.) *Keyness in Texts*. John Benjamins. 21-42.

Stucky, N. and Wimmer, C. (2002). (eds.) *Teaching Performance Studies*. Carbondale & Edwardsville: Southern Illinois University Press.

Suthers, D. D. (2005). Technology affordances for intersubjective learning: A thematic agenda for CSCL. In *Proceedings of th 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!* (pp. 662-671). International Society of the Learning Sciences.

Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Sweetser, E. (1990). *From Etymology to Pragmatics: The Mind-body Metaphor in Semantic Structure and Semantic Change*. Cambridge: CUP.

Sweetser, E. (1999). 'Compositionality and blending: semantic composition in a cognitively realistic framework' *Cognitive Linguistics: Foundations, Scope, and Methodology*, 129-162.

Sweetser, E. and Fauconnier, G. (1996). 'Cognitive links and domains: Basic aspects of Mental Space Theory'. In Fauconnier, G. and Sweetser, E. (eds.) *Spaces, Worlds and Grammars*. Chicago; London: University of Chicago Press. 1-28.

Syh-Jong, J. (2007). “A study of students' construction of science knowledge: Talk and writing in a collaborative group”. *Educational Research*, 49(1), 65-81.

Tannen, D. (1993). ‘What’s in a frame? Surface evidence for underlying expectations’. In Tannen, D. (ed.), *Framing in Discourse*. New York: Oxford University Press. 14–56.

Tannen, D. and Wallat, C. (1986). ‘Medical professionals and parents: A linguistic analysis of communication across contexts’. *Language in Society*, 15(3), 295-311.

Tannen, D. and Wallat, C. (1987). ‘Interactive frames and knowledge schemas in interaction: Examples from a medical examination/interview’. *Social Psychology Quarterly*, 50(2), 205-216.

Tannen, D. and Wallat, C. (1993). ‘Interactive frames and knowledge schemas in interaction: examples from a medical examination/interview’. In Tannen, D. (ed.), *Framing in Discourse*. Oxford: Oxford University Press. 57-76.

Telles-Ribeiro, B., and Hoyle, S. M. (2009). ‘Frame analysis’. In F. Brisard, J.-O. Ostman, and J. Verschueren (eds.), *Grammar, Meaning and Pragmatics*. Amsterdam: John Benjamins. 74-90.

ten Have, P. (2007). *Doing Conversation Analysis*. London: Sage Publications.

Terasaki, A. K. (2004). ‘Pre-announcement sequences in conversation’. In Lerner, G. (ed.) *Conversation Analysis: Studies from the First Generation*. Amsterdam: John Benjamins. 171-223.

Teubert, W. (2007). ‘Language and corpus linguistic’. In W. Teubert & A. Cermakova (eds.), *Corpus Linguistics: A Short Introduction*. London; New York: Continuum. 1-58.

Tomasello, M. (1995). ‘Joint attention as social cognition’. In Moore, C. and Dunham, P. (eds.) *Joint Attention: Its Origins and Role in Development*. New York; London: Psychology Press. Taylor & Francis Group. 103-130.

Tottie, G. (1991). ‘Conversational Style in British and American English: The case of backchannels’. In: Aijmer, K. and Altenberg, B. (eds) *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. New York: Longman. 254-271.

Tracy, K. (1997). *Colloquium: Dilemmas of Academic Discourse*. Westport, CT.: Ablex Publishing.

Tribble, A. and Fahlman, S.E., (2007). ‘CMU-AT: Semantic distance and background knowledge for identifying semantic relations’. In *Proceedings of the 4th International Workshop on Semantic Evaluations* (pp. 121-124). Association for Computational Linguistics.

Tribble, C. (2000). ‘Genres, keywords, teaching: towards a pedagogic account of the language of project proposals.’ In Burnard, L. and McEnery, T. (Eds.), *Rethinking Language Pedagogy from a Corpus Perspective: Papers from the third International Conference on Teaching and Language Corpora*. Frankfurt: Peter Lang. 75–90.

Turner, M. (1986). ‘An image-schematic constraint on metaphor’. In Geiger, R. A. and Rudzka-Ostyn, B. (eds.) *Conceptualizations and Mental Processing in Language*. Mouton de Gruyter. 291-306.

Turner, M. (1993). *Reading Minds: The Study of English in the Age of Cognitive Science*. Princeton, New Jersey: Princeton University Press.

Tyler, A. and Evans, V. (2003). *The Semantics of English Prepositions: Spatial Scenes, Embodied Meaning, and Cognition*. Cambridge: Cambridge University Press.

Uccelli, P., Barr, C.D., Dobbs, C.L., Galloway, E. P., Meneses, A. and Sanchez, E. (2014). “Core academic language skills: An expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners”. *Applied Psycholinguistics*, 36(5), 1077-1109.

Uhmann, S. (1996). On rhythm in everyday German conversation: beat clashes in assessment utterances. In E. Couper-Kuhlen, & M. Selting (eds.), *Prosody in Conversation: Interactional Studies*. Cambridge: Cambridge University Press. 303-365.

van Aalst, J. (2009). 'Distinguishing knowledge-sharing, knowledge-construction, and knowledge-creation discourses'. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 259-287.

van Boxtel, C., van der Linden, J. and Kanselaar, G. (2000). 'Collaborative learning tasks and the elaboration of conceptual knowledge'. *Learning and Instruction*, 10, 311-330.

van Kuppevelt, J. (1995). Discourse structure, topicality and questioning. *Journal of Linguistics*, 31(1), 109-147.

Viana, V., Zyngier, S. and Barnbrook, G. (2011). *Perspectives on Corpus Linguistics*. Amsterdam: John Benjamins Publishing Company.

Viechnicki, G. B. (1997). 'An empirical analysis of participant intentions: Discourse in a graduate seminar'. *Language & Communication*, 17(2), 103-131.

von Glasersfeld, E. (1995). *Radical Constructivism: A Way of Knowing and Doing.* London: Falmer.

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological processes*. Cambridge Mass: Harvard University Press.

Walsh, S. (2002). 'Construction or obstruction: teacher talk and learner involvement in the EFL classroom'. *Language Teaching Research*, 6(1), 3-23.

Walsh, S. (2006). *Investigating Classroom Discourse*. London: Routledge.

Walsh, S. (2014). Outline of NUCASE update and planning.

Walsh, S. and O'Keeffe, A. (2007). 'Applying CA to a modes analysis of third-level spoken academic discourse'. In H. Boyles, and P. Seedhouse (eds.) *Conversation Analysis in Languages for Specific Purpose*. Cambridge: Cambridge University Press. 101-139.

Walsh, S., Morton, T., and O'Keeffe, A. (2011). 'Analyzing university spoken interaction: A CL/CA approach'. *International Journal of Corpus Linguistics*, 16(3), 326-345.

Walsh, S. and Knight, D. (2016). 'Analysing spoken discourse in university small group teaching'. In: *Creating and Digitizing Language Corpora: Volume 3: Databases for Public Engagement*. Palgrave Macmillan. 291-319.

Walsh, S. and Li, L. (2013). 'Conversations as space for learning'. *International Journal of Applied Linguistics*. 23(2), 247-266.

Walsh, S. and Li, L. (2016). 'Classroom talk, interaction and collaboration'. In: G. Hall, ed. *Routledge Handbook of English Language Teaching*. Routledge.

Wang, Q., Woo, H. L. and Zhao, J. (2009). 'Investigating critical thinking and knowledge construction in an interactive learning environment'. *Interactive Learning Environments*, 17(1), 95-104.

Waxman, S.R., & D.G. Hall. (1993). 'The development of a linkage between count nouns and object categories: evidence from fifteen- to twenty-month-old infants'. *Child Development*, 64, 1224-1241.

Weatherall, A. (2011). '*I don't know* as a prepositioned epistemic hedge'. *Research on Language and Social Interaction*, 44 (4), 317-337.

Webb, N.M. (1989). 'Peer interaction and learning in small groups'. *International Journal of Educational Research*, 13, 21-39.

Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. NY: Cambridge.

Westgate, D. and Corden, R. (1993). 'What we thought about things': Expectations, context and small-group talk'. *Language and Education*, 7(2), 115-126.

Whitsitt, S. (2005). 'A critique of the concept of semantic prosody'. *International Journal of Corpus Linguistics*, 10(3), 283-305.

Widdowson, H. G. (1998). 'Communication and community: The pragmatics of ESP'. *English for Specific Purposes*, 17(1), 3-14.

Widdowson, H. G. (2000) 'On the Limitations of Linguistics Applied'. *Applied Linguistics*, 21(1), 3–25.

Williams, G. (1998). 'Collocational networks: Interlocking patterns of lexis in a corpus of plant biology research articles'. *International Journal of Corpus Linguistics, 3*(1), 151-171.

Wolff, P. and Holmes, K. J. (2011). 'Linguistic Relativity'. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(3), 253-265.

Wolfsdorf, D. (2011). 'Plato's conception of knowledge'. *The Classical World*, 105(1), 57-75.

Woods, A., Fletcher, P. and Hughes, A. (1986). *Statistics in Language Studies*. Cambridge: Cambridge University Press.

Wooten, D. B. and Reed, A. (2000). A conceptual overview of the self-presentational concern and response tendencies of focus group participants. *Journal of Consumer Psychology*, 9, 142-153.

Wray, A. and Perkins, M. R. (2000). The functions of formulaic language: an integrated model. *Language & Communication*, 20. 1–28.

Xiao, R. and McEnery, T. (2006). 'Collocation, semantic prosody, and near synomyny: A corpus-linguistic perspective'. *Applied linguistics, 27*(1), 103 ‑ 129.

Yates, L. (2003). ‘Interpretive claims and methodological warrant in small-number qualitative, longitudinal research’. *International Journal of Social Research Methodology*, 6(3), 223-232.

Young, R. F. (2000). ‘Interactional Competence: Challenges for validity’. Paper presented at a joint symposium on “Interdisciplinary Interfaces with Language Testing”, the annual meeting of *the American Association for Applied Linguistics (AAAL)*, Vancouver, British Columbia, Canada.

Young, R. F. (2011). ‘Interactional competence in language learning, teaching, and testing’. *Handbook of Research in Second Language Teaching and Learning*, 2, 426-443.

Young, R. F. (2013). ‘Learning to talk the talk and walk the walk: Interactional competence in academic spoken English’. *Ibérica*, 25, 15-38.

Yule, G. (1996). *Pragmatics*. Oxford: Oxford University Press.

Appendices

Appendix A. USAS Semantic Tagset

USAS Semantic Tagset

See http://ucrel.lancs.ac.uk/usas/ for more details.

A GENERAL & ABSTRACT TERMS	
A1	General
A1.1.1	General actions, making etc.
A1.1.2	Damaging and destroying
A1.2	Suitability
A1.3	Caution
A1.4	Chance, luck
A1.5	Use
A1.5.1	Using
A1.5.2	Usefulness
A1.6	Physical/mental
A1.7	Constraint
A1.8	Inclusion/Exclusion
A1.9	Avoiding
A2	Affect
A2.1	Affect: Modify, change
A2.2	Affect: Cause/Connected
A3	Being
A4	Classification
A4.1	Generally kinds, groups, examples
A4.2	Particular/general; detail
A5	Evaluation
A5.1	Evaluation: Good/bad
A5.2	Evaluation: True/false
A5.3	Evaluation: Accuracy
A5.4	Evaluation: Authenticity
A6	Comparing
A6.1	Comparing: Similar/different
A6.2	Comparing: Usual/unusual
A6.3	Comparing: Variety
A7	Definite (+ modals)
A8	Seem
A9	Getting and giving; possession
A10	Open/closed; Hiding/Hidden; Finding; Showing
A11	Importance
A11.1	Importance: Important
A11.2	Importance: Noticeability
A12	Easy/difficult
A13	Degree
A13.1	Degree: Non-specific
A13.2	Degree: Maximizers
A13.3	Degree: Boosters
A13.4	Degree: Approximators
A13.5	Degree: Compromisers
A13.6	Degree: Diminishers
A13.7	Degree: Minimizers
A14	Exclusivizers/particularizers
A15	Safety/Danger
B THE BODY & THE INDIVIDUAL	
B1	Anatomy and physiology
B2	Health and disease
B3	Medicines and medical treatment
B4	Cleaning and personal care
B5	Clothes and personal belongings
C ARTS & CRAFTS	
C1	Arts and crafts
E EMOTIONAL ACTIONS, STATES & PROCESSES	
E1	General
E2	Liking
E3	Calm/Violent/Angry
E4	Happy/sad
E4.1	Happy/sad: Happy
E4.2	Happy/sad: Contentment
E5	Fear/bravery/shock
E6	**Worry, concern, confident**
F FOOD & FARMING	
F1	Food
F2	Drinks
F3	Cigarettes and drugs
F4	Farming & Horticulture
G GOVT. & THE PUBLIC DOMAIN	
G1	Government, Politics & elections
G1.1	Government etc.
G1.2	Politics
G2	Crime, law and order
G2.1	Crime, law and order: Law & order
G2.2	General ethics
G3	Warfare, defence and the army; Weapons
H ARCHITECTURE, BUILDINGS, HOUSES & THE HOME	
H1	Architecture, kinds of houses & buildings
H2	Parts of buildings
H3	Areas around or near houses
H4	Residence
H5	Furniture and household fittings
I MONEY & COMMERCE	
I1	Money generally
I1.1	Money: Affluence
I1.2	Money: Debts
I1.3	Money: Price
I2	Business
I2.1	Business: Generally
I2.2	Business: Selling
I3	Work and employment
I3.1	Work and employment: Generally
I3.2	Work and employment: Professionalism
I4	Industry
K ENTERTAINMENT, SPORTS & GAMES	
K1	Entertainment generally
K2	Music and related activities
K3	Recorded sound etc.
K4	Drama, the theatre & show business
K5	Sports and games generally
K5.1	Sports
K5.2	Games
K6	Children's games and toys
L LIFE & LIVING THINGS	
L1	Life and living things
L2	Living creatures generally
L3	Plants
M MOVEMENT, LOCATION, TRAVEL & TRANSPORT	
M1	Moving, coming and going
M2	Putting, taking, pulling, pushing, transporting &c.
M3	Movement/transportation: land
M4	Movement/transportation: water
M5	Movement/transportation: air
M6	Location and direction
M7	Places
M8	Remaining/stationary
N NUMBERS & MEASUREMENT	
N1	Numbers
N2	Mathematics
N3	Measurement
N3.1	Measurement: General
N3.2	Measurement: Size
N3.3	Measurement: Distance
N3.4	Measurement: Volume
N3.5	Measurement: Weight
N3.6	Measurement: Area
N3.7	Measurement: Length & height
N3.8	Measurement: Speed
N4	Linear order
N5	Quantities
N5.1	Entirety; maximum
N5.2	Exceeding; waste
N6	Frequency etc.
O SUBSTANCES, MATERIALS, OBJECTS & EQUIPMENT	
O1	Substances and materials generally
O1.1	Substances and materials generally: Solid
O1.2	Substances and materials generally: Liquid
O1.3	Substances and materials generally: Gas
O2	Objects generally
O3	Electricity and electrical equipment
O4	Physical attributes
O4.1	General appearance and physical properties
O4.2	Judgement of appearance (pretty etc.)
O4.3	Colour and colour patterns
O4.4	Shape
O4.5	Texture
O4.6	Temperature
P EDUCATION	
P1	Education in general
Q LINGUISTIC ACTIONS, STATES & PROCESSES	
Q1	Communication
Q1.1	Communication in general
Q1.2	Paper documents and writing
Q1.3	Telecommunications
Q2	Speech acts
Q2.1	Speech etc: Communicative
Q2.2	Speech acts
Q3	Language, speech and grammar
Q4	The Media
Q4.1	The Media: Books
Q4.2	The Media: Newspapers etc.
Q4.3	The Media: TV, Radio & Cinema
S SOCIAL ACTIONS, STATES & PROCESSES	
S1	Social actions, states & processes
S1.1	Social actions, states & processes
S1.1.1	General
S1.1.2	Reciprocity
S1.1.3	Participation
S1.1.4	Deserve etc.
S1.2	Personality traits
S1.2.1	Approachability and Friendliness
S1.2.2	Avarice
S1.2.3	Egoism
S1.2.4	Politeness
S1.2.5	Toughness; strong/weak
S1.2.6	Sensible
S2	People
S2.1	People: Female
S2.2	People: Male
S3	Relationship
S3.1	Relationship: General
S3.2	Relationship: Intimate/sexual
S4	Kin
S5	Groups and affiliation
S6	Obligation and necessity
S7	Power relationship
S7.1	Power, organizing
S7.2	Respect
S7.3	Competition
S7.4	Permission
S8	Helping/hindering
S9	Religion and the supernatural
T TIME	
T1	Time
T1.1	Time: General
T1.1.1	Time: General: Past
T1.1.2	Time: General: Present; simultaneous
T1.1.3	Time: General: Future
T1.2	Time: Momentary
T1.3	Time: Period
T2	Time: Beginning and ending
T3	Time: Old, new and young; age
T4	Time: Early/late
W THE WORLD & OUR ENVIRONMENT	
W1	The universe
W2	Light
W3	Geographical terms
W4	Weather
W5	Green issues
X PSYCHOLOGICAL ACTIONS, STATES & PROCESSES	
X1	General
X2	Mental actions and processes
X2.1	Thought, belief
X2.2	Knowledge
X2.3	Learn
X2.4	Investigate, examine, test, search
X2.5	Understand
X2.6	Expect
X3	Sensory
X3.1	Sensory: Taste
X3.2	Sensory: Sound
X3.3	Sensory: Touch
X3.4	Sensory: Sight
X3.5	Sensory: Smell
X4	Mental object
X4.1	Mental object: Conceptual object
X4.2	Mental object: Means, method
X5	Attention
X5.1	Attention
X5.2	Interest/boredom/excited/energetic
X6	Deciding
X7	Wanting; planning; choosing
X8	Trying
X9	Ability
X9.1	Ability: Ability, intelligence
X9.2	Ability: Success and failure
Y SCIENCE & TECHNOLOGY	
Y1	Science and technology in general
Y2	Information technology and computing
Z NAMES & GRAMMATICAL WORDS	
Z0	Unmatched proper noun
Z1	Personal names
Z2	Geographical names
Z3	Other proper names
Z4	Discourse Bin
Z5	Grammatical bin
Z6	Negative
Z7	If
Z8	Pronouns etc.
Z9	Trash can
Z99	Unmatched

Appendix B. Full lists of over-used key semantic domains across the sub corpora

Ranking	Item	O1	1%	O2	2%	Over/under-use	LL	Log Ratio	Description
1	P1	2109	1.41	16329	0.48	+	1643.65	1.55	Education in general
2	Z8	26828	17.92	499076	14.71	+	937.93	0.29	Pronouns
3	E4.1+	457	0.31	1548	0.05	+	873.24	2.74	Happy
4	Z4	9021	6.03	152983	4.51	+	659.02	0.42	Discourse Bin
5	Q2.1	2526	1.69	32756	0.97	+	625.83	0.81	Speech: Communicative
6	E6+	116	0.08	362	0.01	+	235.57	2.86	Confident
7	X2.1	1648	1.1	25570	0.75	+	199.48	0.55	Thought, belief
8	S1.1.2+	210	0.14	1552	0.05	+	175.64	1.62	Reciprocal
9	Z6	2869	1.92	50387	1.48	+	163.99	0.37	Negative
10	A13.5	270	0.18	2448	0.07	+	160.76	1.32	Degree: Compromisers
11	A5.4+	524	0.35	6342	0.19	+	160.27	0.91	Evaluation: Authentic
12	X3.2-	73	0.05	244	0.01	+	140.91	2.76	Sound: Quiet
13	A14	1278	0.85	20464	0.6	+	131.42	0.5	Exclusivizers/particularizers
14	X2.2+	1407	0.94	23019	0.68	+	128.26	0.47	Knowledgeable
15	X2.3+	179	0.12	1487	0.04	+	124.48	1.45	Learning
16	S1.1.3+	134	0.09	936	0.03	+	121.54	1.7	Participating
17	X3.2	228	0.15	2372	0.07	+	102.38	1.12	Sensory: Sound
18	X8+	348	0.23	4378	0.13	+	94.9	0.85	Trying hard
19	N4	1539	1.03	26860	0.79	+	92.15	0.38	Linear order
20	A7+	2649	1.77	49406	1.46	+	90.29	0.28	Likely
21	X4.1	639	0.43	9645	0.28	+	88.2	0.59	Mental object: Conceptual object
22	A1.1.1	2243	1.5	41892	1.23	+	75.4	0.28	General actions / making
23	I3.2+	55	0.04	340	0.01	+	58.58	1.87	Professional
24	E6-	170	0.11	2014	0.06	+	55.22	0.94	Worry
25	Q2.2	1351	0.9	24771	0.73	+	54.14	0.31	Speech acts
26	A7+++	88	0.06	800	0.02	+	52.13	1.32	Likely
27	E2+	419	0.28	6496	0.19	+	50.92	0.55	Like
28	A3+	7747	5.17	161836	4.77	+	48.1	0.12	Existing
29	A1.8+	185	0.12	2410	0.07	+	45.15	0.8	Inclusion
30	A13.6	215	0.14	2976	0.09	+	42.44	0.71	Degree: Diminishers
31	X2.2	97	0.06	1026	0.03	+	41.94	1.1	Knowledge
32	A1.2+	66	0.04	598	0.02	+	39.35	1.32	Suitable
33	A7	284	0.19	4320	0.13	+	37.82	0.58	Probability
34	I3.1	404	0.27	6676	0.2	+	34.72	0.46	Work and employment: Generally
35	Q1.3	87	0.06	1032	0.03	+	28.16	0.93	Telecommunications
36	N5-	269	0.18	4348	0.13	+	26.27	0.49	Quantities: little
37	Z7	980	0.65	18822	0.55	+	24.35	0.24	If
38	X9.1+	172	0.11	2598	0.08	+	23.66	0.59	Able/intelligent
39	A1.5.2+	45	0.03	450	0.01	+	22.03	1.18	Useful
40	B2	32	0.02	276	0.01	+	20.86	1.39	Health and disease
41	X5.2+	273	0.18	4620	0.14	+	20.22	0.42	Interested/excited/energetic
42	S6+	906	0.61	17591	0.52	+	19.73	0.22	Strong obligation or necessity
43	X9.1	50	0.03	550	0.02	+	19.7	1.04	Ability and intelligence
44	N6	99	0.07	1398	0.04	+	18.11	0.68	Frequency

Table B1: Over-used key semantic domains in ALE sub corpus

Ranking	Item	O1	1%	O2	2%	Over/und	LL	Log Ratio	Description
1	Y2	903	0.64	3666	0.11	+	1566.77	2.56	Information technology and computing
2	E4.1+	507	0.36	1548	0.05	+	1092.16	2.97	Happy
3	Z4	9079	6.41	152983	4.51	+	955.57	0.51	Discourse Bin
4	Z8	25300	17.85	499076	14.71	+	853.99	0.28	Pronouns
5	A14	1733	1.22	20464	0.6	+	657.88	1.02	Exclusivizers/particularizers
6	T2-	545	0.38	3848	0.11	+	526.86	1.76	Time: Ending
7	A1.1.1	2751	1.94	41892	1.23	+	464.65	0.65	General actions / making
8	Q1.2	952	0.67	12035	0.35	+	301.52	0.92	Paper documents and writing
9	S6+	1254	0.88	17591	0.52	+	287.79	0.77	Strong obligation or necessity
10	X3.2-	106	0.07	244	0.01	+	272.62	3.38	Sound: Quiet
11	X8+	413	0.29	4378	0.13	+	201.34	1.18	Trying hard
12	I3.1	536	0.38	6676	0.2	+	176.91	0.94	Work and employment: Generally
13	X2	133	0.09	728	0.02	+	174.09	2.13	Mental actions and processes
14	S3.2	211	0.15	1674	0.05	+	172.87	1.59	Relationship: Intimacy and sex
15	T1.1.3	1421	1	23692	0.7	+	158.15	0.52	Time: Future
16	F1	365	0.26	4224	0.12	+	145.65	1.05	Food
17	A7-	108	0.08	684	0.02	+	119.86	1.92	Unlikely
18	M2	784	0.55	12663	0.37	+	102.15	0.57	Putting, pulling, pushing, transporting
19	X3.2	210	0.15	2372	0.07	+	88.8	1.08	Sensory: Sound
20	B3	268	0.19	3446	0.1	+	80.87	0.9	Medicines and medical treatment
21	X2.2	116	0.08	1026	0.03	+	79.86	1.44	Knowledge
22	F2	143	0.1	1452	0.04	+	76.24	1.24	Drinks and alcohol
23	A10+	499	0.35	7792	0.23	+	75.81	0.62	Open; Finding; Showing
24	Z6	2516	1.78	50387	1.48	+	72.57	0.26	Negative
25	A1.5.1	432	0.3	6604	0.19	+	71.81	0.65	Using
26	E2+	424	0.3	6496	0.19	+	69.84	0.64	Like
27	Q1.3	107	0.08	1032	0.03	+	63.07	1.31	Telecommunications
28	O4.2+	200	0.14	2554	0.08	+	61.56	0.91	Judgement of appearance: Beautiful
29	T1.1.1	377	0.27	5816	0.17	+	60.33	0.63	Time: Past
30	A5.4+	403	0.28	6342	0.19	+	59.17	0.61	Evaluation: Authentic
31	M4	106	0.07	1062	0.03	+	58.05	1.26	Sailing, swimming, etc.
32	Q2.1-	16	0.01	18	0	+	57.39	4.41	Speech: Not communicating
33	M1	1367	0.96	26957	0.79	+	46.29	0.28	Moving, coming and going
34	O2	1173	0.83	22784	0.67	+	45.92	0.3	Objects generally
35	S1.2.6-	55	0.04	446	0.01	+	43.58	1.56	Foolish
36	X3.3	39	0.03	268	0.01	+	39.07	1.8	Sensory: Touch
37	N3.5	57	0.04	570	0.02	+	31.33	1.26	Measurement: Weight
38	X2.4	311	0.22	5470	0.16	+	25.59	0.45	Investigate, examine, test, search
39	A9	26	0.02	244	0.01	+	16.13	1.35	Getting and giving; possession
40	A1.5.2-	12	0.01	66	0	+	15.63	2.12	Useless
41	S2.2	232	0.16	4216	0.12	+	15.43	0.4	People: Male

Table B2: Over-used key semantic domains in BI sub corpus

Ranking	Item	O1	1%	O2	2%	Over/und	LL	Log Ratio	Description
1	M4	969	0.36	1062	0.03	+	2393.58	3.51	Sailing, swimming, etc.
2	W4	563	0.21	758	0.02	+	1243.42	3.21	Weather
3	I1.3--	24	0.01	42	0	+	44.81	2.83	Cheap
4	O3	643	0.24	1146	0.03	+	1185.1	2.81	Electricity and electrical equipment
5	N3.3	492	0.18	940	0.03	+	861.87	2.71	Measurement: Distance
6	E4.1+	747	0.27	1548	0.05	+	1228.85	2.59	Happy
7	N3.5	269	0.1	570	0.02	+	434.66	2.56	Measurement: Weight
8	S1.1.3+	413	0.15	936	0.03	+	630.96	2.46	Participating
9	A1.5.2-	22	0.01	66	0	+	25.66	2.06	Useless
10	I1.3+	52	0.02	180	0.01	+	51.38	1.85	Expensive
11	W3	1035	0.38	4486	0.13	+	748	1.53	Geographical terms
12	X6	91	0.03	412	0.01	+	61.32	1.46	Deciding
13	X3.3	59	0.02	268	0.01	+	39.54	1.46	Sensory: Touch
14	O4.6+	246	0.09	1164	0.03	+	153.86	1.4	Temperature: Hot / on fire
15	N5.1+++	28	0.01	132	0	+	17.62	1.4	Entire; maximum
16	A15+	62	0.02	298	0.01	+	37.73	1.38	Safe
17	I1.3	334	0.12	1618	0.05	+	200.43	1.37	Money: Cost and price
18	G2.2+	166	0.06	800	0.02	+	100.54	1.37	Ethical
19	T1.2	391	0.14	1910	0.06	+	231.15	1.35	Time: Momentary
20	O4.6-	182	0.07	962	0.03	+	92.64	1.24	Temperature: Cold
21	N3.8+	380	0.14	2056	0.06	+	184.46	1.21	Speed: Fast
22	E6+	66	0.02	362	0.01	+	31.14	1.19	Confident
23	A5.1+++	261	0.1	1554	0.05	+	102.54	1.07	Evaluation: Good
24	A12++	55	0.02	326	0.01	+	21.84	1.07	Easy
25	O4.1	897	0.33	5456	0.16	+	334.76	1.04	General appearance and physical properties
26	A13.5	395	0.15	2448	0.07	+	140.71	1.01	Degree: Compromisers
27	X6+	267	0.1	1664	0.05	+	93.77	1	Decided
28	Q1.3	159	0.06	1032	0.03	+	50.18	0.94	Telecommunications
29	T1.1.3	3610	1.33	23692	0.7	+	1105.11	0.93	Time: Future
30	C1	804	0.3	5410	0.16	+	229.28	0.89	Arts and crafts
31	O1.2	297	0.11	1996	0.06	+	85	0.89	Substances and materials: Liquid
32	N3.2	165	0.06	1116	0.03	+	46.35	0.88	Measurement: Size
33	A5.4+	898	0.33	6342	0.19	+	221.13	0.82	Evaluation: Authentic
34	I4	105	0.04	744	0.02	+	25.59	0.82	Industry
35	X2.4	768	0.28	5470	0.16	+	184.06	0.81	Investigate, examine, test, search
36	I2.1	306	0.11	2210	0.07	+	70.07	0.79	Business: Generally
37	N3.3+	108	0.04	779	0.02	+	24.84	0.79	Distance: Far
38	I3.1	917	0.34	6676	0.2	+	204.44	0.78	Work and employment: Generally
39	O2	3073	1.13	22784	0.67	+	643.5	0.75	Objects generally
40	S6+	2337	0.86	17591	0.52	+	463.75	0.73	Strong obligation or necessity
41	Z4	19808	7.28	152983	4.51	+	3572.05	0.69	Discourse Bin
42	A7+++	103	0.04	800	0.02	+	18.18	0.68	Likely
43	N5-	554	0.2	4348	0.13	+	93.84	0.67	Quantities: little
44	A14	2588	0.95	20464	0.6	+	425.45	0.66	Exclusivizers/particularizers
45	M2	1596	0.59	12663	0.37	+	258.77	0.65	Putting, pulling, pushing, transporting
46	A13.6	368	0.14	2976	0.09	+	55.11	0.63	Degree: Diminishers
47	A5.4-	105	0.04	846	0.02	+	15.97	0.63	Evaluation: Unauthentic
48	T2+	849	0.31	6936	0.2	+	121.67	0.61	Time: Beginning
49	A1.1.1	4616	1.7	41892	1.23	+	387.72	0.46	General actions / making
50	T1.3	1690	0.62	15325	0.45	+	142.63	0.46	Time: Period
51	A13.2	473	0.17	4394	0.13	+	34.4	0.43	Degree: Maximizers
52	M3	325	0.12	3048	0.09	+	22.21	0.41	Vehicles and transport on land
53	N5++	1107	0.41	10539	0.31	+	68.27	0.39	Quantities: many/much
54	S7.1+	763	0.28	7283	0.21	+	46.19	0.39	In power
55	N5.1+	1814	0.67	17791	0.52	+	89.57	0.35	Entire; maximum
56	X7+	1709	0.63	17022	0.5	+	74.32	0.33	Wanted
57	Z7	1868	0.69	18822	0.55	+	73.41	0.31	If
58	A13.4	493	0.18	5040	0.15	+	16.93	0.29	Degree: Approximators
59	M1	2608	0.96	26957	0.79	+	80.24	0.27	Moving, coming and going
60	N4	2586	0.95	26860	0.79	+	75.63	0.26	Linear order
61	A7+	4630	1.7	49406	1.46	+	99.47	0.23	Likely
62	Q1.2	1130	0.42	12035	0.35	+	24.86	0.23	Paper documents and writing
63	Z8	45989	16.91	499076	14.71	+	791.58	0.2	Pronouns
64	N5+	1199	0.44	13019	0.38	+	20.48	0.2	Quantities: many/much
65	A5.1+	3632	1.34	40101	1.18	+	48.33	0.18	Evaluation: Good

Table B3: Over-used key semantic domains in ME sub corpus

Appendix C. A partial list of Jeffersonian transcribing conventions

(.) A full stop inside brackets denotes a micro pause, a notable pause but of no significant length.

(0.2) A number inside brackets denotes a timed pause. This is a pause long enough to time and subsequently show in transcription.

[Square brackets denote a point where overlapping speech occurs.

> < Arrows surrounding talk like these show that the pace of the speech has quickened

< > Arrows in this direction show that the pace of the speech has slowed down

( ) Where there is space between brackets denotes that the words spoken here were too unclear to transcribe

(( )) Where double brackets appear with a description inserted denotes some contextual information where no symbol of representation was available.

Underline When a word or part of a word is underlined it denotes a raise in volume or emphasis

↑ When an upward arrow appears it means there is a rise in intonation

↓ When a downward arrow appears it means there is a drop in intonation

→ An arrow like this denotes a particular sentence of interest to the analyst

CAPITALS where capital letters appear it denotes that something was said loudly or even shouted

Hum(h)our When a bracketed 'h' appears it means that there was laughter within the talk

= The equal sign represents latched speech, a continuation of talk

:: Colons appear to represent elongated speech, a stretched sound

°example° low volume
°°example°°much lower volume (e.g. whispering)

Appendix D. Excerpt 16

Excerpt 16. ***propeller design***

```
1   <$1>  Was that you= you design your propeller ↓first
2   <$7>  Yeah
3   <$1>  (.) and choose your generator from your propeller
4         ↑design(.)You're ↓no:t (.)choose your generator
5         first(.)and then
6         [choose your propeller design(.) yeah your right
7   <$3>  [This is the thing as well(.) ↑CAN YOU: em::-
8   <$1>  =°Is this aimed at ↑me or ((anonymous names)) ↓Me°
9   <$3>  Yeah(.) it's you(.) Em::(1.0) I'm just thinking(1.4) <can
10        you> predict how::(1.3) if you had a propeller ↑design in
11        these seas how fast it would= how: fast it would ro↓tate
12  <$1>  Yeah
13  <$3>  Right(.) I'm just trying to think if it's rota:ting
14        (1.2)on a ((yawning))generator(.)it requires
15        torque(.)doesn't ↑it(.) So would you be able to calculate
16        how much tor:que(.)[it would be producing
17  <$6>                     [That's- that's where I'm going to
18        come ↑in(.) I wanna do the ↑modeling of ↑all of ↑that-
19  <$3>  =Right(.) I'm yeah= I'm just trying to think is(.) what-
20        <how: do you wor:k it out first(.) can you wor:k out>
21  <$1>  (0.8)Well you do it-
22  <$3>  =The tu= the TURBINE SPINS(.) can you work out how much
23        tor:que that is producing
24  <$1>  =when you did your ↑propeller design(.)[I think but-
25  <$3>                                        [No I'm just
26        trying to think(.) you do= so you don't really need a
27        gene↑rator(.)there
28  <$1>  (1.2)Eh::(.) No because
29        [the generator will have some resistance= won't it
30  <$6>  [Well you see(.) the- well- [The re= yeah(.) >But the was
31        obviously<-
32  <$3>                              [Yeah(.) >That's the= that's
33        the= re= that's what we're trying to say is<(.) Yeah
34  <$6>  =we want it to= we want it to= be completely compatible
35        so what I wha= this is my inten↑tion for the project(.)
36        Let's say we're doing this and we're testing and (as) it
37        goes(.) I want to(.) <li:nk how that propeller is
38        behaving>(.) ↓to the electrical(.) to a= ↓to system on-
39  <$1>  Yeah
40  <$6>  =the computer(.) and >to kind of< ↑model ↑it(.)and then
41        you'll see what-
42  <$1>  Yeah
43  <$6>  =You ↑know(.) what ↑outputs you're getting what ↑torque
44        you're getting what ↑currents you're ↓getting
45  <$1>  Yeah
46  <$6>  And then- well that's the intention(.) so we a= so it's
47        an actual ↑compatibility
```

Appendix E. Metadata of small group talk scenarios

Scenario No.	Lexical item in focus	Main topic	Cognitive frame (FrameNet)	File name	SGT type	No. of speakers	Tutor code	Time stamp
1	*student-teacher*	Student-teachers in the work place	[IDENTITY]; [BOCOMING_DETACHED]	NC023	Tutored group meeting	5	<$1>, <$2>	42s (00:09:32 – 00:10:14)
2	*student-teacher*	Institutional expectation of student-teachers	[EXPECTATION]	NC023	Tutored group meeting	5	<$1>, <$2>	01m:38s (01:02:02 – 01:04:01)
3	*academic*	Discourse markers in academic talk	[EDUCATION_TEACHING]	NC004	Seminar	1	<$1>	01m:42s (01:45:34 – 01:47:16)
4	*transcript*	Non-verbal features in transcripts	[INFORMATION_DISPLAY]	NC073	Tutored group discussion	4	<$4>, <$5>	03m:15s (00:27:07 – 00:30:22)
5	*transcript*	Adjacency pairs in transcripts	[SCRUTINY]	NC026	Tutored group discussion	4	<$2>,<3>, <$4>	02m:41s (01:00:45 – 01:03:26)
6	*Social conscience*	Social conscience	[PERCEPTION_EXPERIENCE]	NC026	Tutored group discussion	5	<$2>,<3>, <$4>	04m:55s (01:10:18 – 01:15:13)
7	*drug*	What to look for in drug targeting	[MEDICAL_INTERVENTION]; [BIOLOGICAL_ENTITY/MECHANISM]; [SIMILARITY]	NC071	Student project meeting	5	<$7>	01m:10s (00:10:27 – 00:11:37)
8	*drug*	Using public database for drug repurposing	[MEDICAL_SPECIALS]	NC071	Student project meeting	4	<$7>	02m:54s (00:03:33 – 00:06:27)

9	*vaccine*	Targeting a vaccine strain	[MEDICAL_INTERVENTION]	NC065	Consultancy meeting	4	N/A	02m:02s (00:01:57 – 00:03:59)
10	*web*	Localizing the web service	Not found in FrameNet	NC142	Student project meeting	3	N/A	N/A
11	*software*	Software development life cycle	[EXECPLAR]	NC075	Requirement gathering meeting	4	N/A	01m:50s (00:00:34 – 00:02:24)
12	*propeller*	Compatibility of the propeller design	[COMPATIBILITY]	NC027	Student group meeting	4	N/A	01m:23s (01:07:42 – 01:09:05)
13	*generator*	The generator getting too hot	[TEMPERATURE]	NC029	Student group meeting	3	N/A	01m:48s (00:20:33 – 00:22:21)
14	*wave*	Reporting wave loading calculation	[STATEMENT]	NC128	Student group meeting	2	N/A	03m:12s (00:02:57 – 00:06:09)
15	*tide*	Sites of tidal farms	[ASSESSING]	NC001	Student group meeting	4	N/A	01m:34s (00:10:24 – 00:11:59)
16	*wind*	Effect of the wind turbine rotating direction	[OBJECTIVE_INFLUENCE]	NC064	Student group meeting	5	N/A	00m:59s (01:04:07 – 01:05:06)

Table E: Metadata of small group talk scenarios