Stance-taking and social status on an online bulletin board

A qualitative and quantitative approach

Nele Põldvere

Centre for Languages and Literature
Master's Thesis 30 hp (SPVR01)
English Language and Linguistics
August 2013

Supervisors:
Carita Paradis
Dylan Glynn

Abstract

In this study, I demonstrate that social hierarchy and power are important aspects for understanding the use of epistemic and evidential stance verbs in computer-mediated communication. The data for the study come from an online bulletin board about rhythmic gymnastics, where the construction of social roles is believed to play a role in the expression of stance. The members of the community are divided into three hierarchically distinct social ranks based on status and activity on the board. I investigate whether members of a higher rank use epistemic and evidential stance verbs in a more authoritative manner than members of lower ranks using two methodological frameworks. In the qualitative part of the study, I adopt the dialogical discourse analysis to argue that epistemic and evidential stance is a dialogically constructed phenomenon that locally emerges between conversational co-participants. The quantitative part of the study employs the multifactorial usage-feature analysis, where two stance verbs *think* and *seem* are coded for a range of formal, semantic and extra-linguistic factors, which are believed to contribute to the differentiation of authoritative and tentative stance. The results show that bulletin board users of a higher rank exhibit a more authoritative and even aggressive use of epistemic and evidential stance verbs than users of lower ranks.

Keywords: corpus linguistics, computer-mediated communication, social hierarchy, dialogical discourse analysis, multifactorial usage-feature analysis

Table of contents

Acknowledgements

My greatest gratitude goes to my supervisors Carita Paradis and Dylan Glynn. First, for letting me work on the data that I had studied in my mind for so long, and secondly, for their invaluable comments and remarks on the earlier drafts of the paper. The whole process has been a truly unforgettable experience.

I am grateful to my friends and family for reviewing parts of this thesis, and my lovely colleagues from the English Linguistics program at Lund University. These two years would not have been the same without you.

I would also like to thank my parents, who accepted the fact that their daughter only came home to write, and my brother, who reserved a table in his office for some serious linguistic research.

List of Figures

List of Tables

Chapter 1

Introduction

This study investigates the socio-cognitive functions of epistemic and evidential stance in Internet bulletin board (rsg.net) conversations. Rsg.net is an asynchronous recreational bulletin board specialized on rhythmic gymnastics. The board displays dialogical and interactive language, where exchange of views and opinions is a prominent feature (Claridge 2007: 97). The approach adopted in this study is based on one of the fundamental views in Cognitive Linguistics, which regards language use as a key factor in structuring grammar and motivating meaning (e.g. Fillmore 1985; Langacker 1987). Studies on the emergent meaning of grammatical and lexical constructions therefore mostly benefit from data sets that display naturally occurring language in conversational settings. For this reason, I have adopted corpus-driven methodological approaches to explore the communicative and conceptual structures of epistemic and evidential stance in bulletin board conversations.

Stance-taking as such is above all an activity through which speakers and writers pursue their socio-functional goals and make sense of the world. Stance is used as a cover term for two types of modality, epistemic and evidential. The definitions of the two modal systems are summarized in Palmer: "[…] with epistemic modality speakers express their judgments about the factual status of the utterance, whereas with evidential modality they indicate the evidence they have for its factual status" ([1986] 2001: 8). Stance does not only represent individual standpoints and judgments, but is framed by values aligned with the various social groups to which speakers and writers belong (Vološinov 1986; Linell 2009). Rsg.net is an example of such a social community, where power and social relations between members of the group are factors that have a strong influence on the flow and nature of interaction. In the present study, stance is treated as a

linguistic phenomenon that highlights social hierarchies between various levels of the group. Consider the following examples:

(1) *I **think** is highly disrespectful as well. How can you say that they were sloppy because they didn't give their best? Maybe this is the current level they are at moment, so putting disgusted faces because they came last is horrible.*
(2) *Naazmi was head and shoulders above everyone else in terms of composition and execution, although the margin between her and second place (17+!) **seems** maybe a touch much.*[1]

Examples (1) and (2) are extracted from rsg.net. The first example is realized by the epistemic mental predicate *think* and the second by the evidential perception verb *seem*. Also, the two examples serve rather different purposes. Example (1) shows the negotiation of power on behalf of an experienced moderator, whose response to a casual user bears a strong marker of authority and assertiveness, as well as commitment to the truth of what he/she is saying: the adjectival phrase *highly disrespectful*. Example (2), on the other hand, is a relatively tentative evaluation of the scores given to a gymnast named Naazmi. Tentativeness here is exemplified by the epistemic adverb *maybe*, followed by another marker of weak commitment, *a touch much*. Both examples are framed either by an epistemic or evidential stance construction and presented in different contextual and constructional environments. The constructions within their immediate contexts demonstrate the multidimensionality of stance-taking in establishing group membership. The users of rsg.net have formed a virtual community that is characterized by social and linguistic variation. In this study, I propose that the expression of stance is a phenomenon that highlights and establishes the social roles these members are assigned to play and their ranks on the board (for the operationalization of social hierarchy in rsg.net, see Chapter 3.1.3).

Stance in the present study is thus restricted to the markers of speakers' commitment to what they are saying (epistemic modality) and the reliability of source on which speakers base their knowledge of the world (evidentiality). Both modalities are grouped under the general term of stance. The lexical items investigated in the study are epistemic and evidential verbs, such as *think, guess, believe, seem, find,* etc. The

[1] All typographic and grammatical errors are preserved in their original form.

complement clauses modified by these verbs are referred to as *propositions*, which function as objects of belief, opinion and other propositional attitudes. These verbs and their propositions are studied in each given context with other grammatical and lexical constructions in compliance with the basic assumptions of Cognitive Linguistics.

Three research questions are central to the purpose of this study. The first research question is an overarching question about the nature of stance constructions used in rsg.net. The following questions address the issue presented in the former, adopting two inherently different methodologies with the aim to reach a more insightful understanding of the functions of epistemic and evidential stance constructions in the data. The questions are as follows:

- Do Internet bulletin board members of a higher rank use epistemic and evidential verbs in a more assertive and authoritative manner than users of a lower rank?
- How do members from three different ranks attenuate or reinforce the strength of their propositions relative to their conversational co-participants?
- What formal, semantic and extra-linguistic factors framed by *think* and *seem* indicate social hierarchy between the three ranks?

Conducting a two-part study on the expression of stance, I will be able to account for the discursive functions of stance constructions in interaction, as well as the multidimensionality of linguistic structuring. To answer the second research question, I conduct a qualitative investigation of extracts from two controversial bulletin board threads. The methodology adopted for the investigation of stance in that section is based on the theoretical assumptions of dialogicality (Linell 1998; Marková 2003), in which interaction is seen as a socially and cognitively constructed phenomenon that emerges intersubjectively between two or more co-participants. The methodological framework used to study the conversational threads in rsg.net is dialogical discourse analysis (Marková et al. 2007; Linell 2009), an approach developed to account for the sequential and thematic organization of text. The method is modified to fit the purposes of the present study, i.e., to investigate the behavior of specific linguistic forms in context. Therefore, the stance constructions found in the two bulletin board threads are studied in interactional discourse units in which they are generated to see how bulletin board users align their epistemic and

evidential verbs with previous and subsequent stance constructions. I attempt to show that power and the lack of power emerge from interaction, and for better understanding of the impact of epistemic and evidential stance on social relations it needs to be studied in larger communicative contexts.

Having studied the dialogical functions of stance-taking, I zoom in on two frequently occurring stance constructions, *think* and *seem*, and adopt a corpus-driven quantitative analysis. Multifactorial usage-feature analysis allows us to quantify the degree to which various formal, semantic and extra-linguistic factors contribute to the meaning of the lexical items under investigation (Geeraerts, Grondelaers, and Bakema 1994; Gries 2003; Divjak 2010; Glynn 2010b). Altogether 729 examples of *think* and *seem* with their linguistic variations are coded for 31 factors with the software Filemaker Pro 11. The examples are then analyzed in the open-source software R using both exploratory (Multiple Correspondence Analysis) and confirmatory statistical tools (Binary Logistic Regression Analysis). This relatively new technique employed in Cognitive Linguistics goes beyond the merely descriptive nature of traditional corpus linguistics, and introduces techniques that allow researchers to answer such complex research questions as the third question presented above.

The present study is organized as follows. Chapter 2 gives an overview of previous works conducted on epistemic and evidential stance, computer-mediated communication and the basic assumptions behind the two methodologies adopted in this study. Social rank is operationalized in Chapter 3, with a detailed description of data collection for both studies. In Chapter 4, I introduce the qualitative method, namely dialogical discourse analysis, and analyze epistemic and evidential stance verbs in their dialogical context. Chapters 5 and 6 both deal with the quantitative part of the study. Chapter 5 introduces the methodology, namely multifactorial usage-feature analysis, and criteria for the coding and analysis of two stance verbs *think* and *seem*. The results are presented in Chapter 6 using both exploratory and confirmatory statistical tools.

Chapter 2
Background

The present chapter gives an overview of the works conducted on stance-taking, computer-mediated communication and the two analytical methods adopted in this study. Section 2.1 examines the representation of stance and, more specifically, epistemic and evidential verbs in previous studies. Section 2.2 is concerned with computer-mediated communication and the issue of power and social hierarchy in virtual communities. Finally, Section 2.3 attempts to introduce the theoretical assumptions behind two methodologies adopted in the present study. To the best of my knowledge, these approaches have not been combined in previous research. However, each of them separately has received considerable amount of attention in literature.

2.1 Previous works on stance-taking

Stance is part of subjective language and subjectivity has been found to be an important component of interaction. Since the second half of the 20th century, linguists have reconsidered the idea of language as a communication of purely propositional and referential material (Jakobson 1960; Lyons 1977). Nevertheless, the insight that human interaction is extensively marked by more or less explicit attitudes towards people, events and concepts in the world is a rather new journey on which linguists have tentatively embarked.

Research in the field of stance markers has shown that stance does not only manifest itself in the face of lexicon, but becomes apparent in the linguistic structures of phonology, syntax and discourse, as well as in the close interaction between them. As a result, such phonological factors as intonation, syntactic constructions as tense and aspect, as well as

conversational features like repetition, all contribute to the idea of language as a 'preeminently subjective' phenomenon (Bybee and Hopper 2001: 7). What is more, recent studies have emphasized the importance of the sociolinguistic element in semantic research (Kristiansen and Dirven 2008; Glynn 2009; Geeraerts, Kristiansen, and Peirsman 2010). Such a broad view to stance has received great interest among researchers working on the phenomenon and established the strong relationship between propositional and evaluative aspects of interaction.

For example, Scheibman (2002) has found, in her large-scale study of linguistic subjectivity, that subjective expressions are predominant in conversational interaction. Scheibman's data comes from audiotaped American English informal conversations, i.e. from naturally occurring language, which contains an extensive usage of subjective markers. Taking into account a considerable number of structural, functional and semantic variables in the construction of stance, Scheibman has managed to observe a large number of structural combinations found in her database. The results imply that speakers' point of view is primarily exemplified by the use of conventionalized subject-predicate constructions (Scheibman 2002: 61). For example, Scheibman finds that *I* and *you* most commonly occur with verbs of cognition or epistemic commitment, including such predicates as *know*, *think*, *guess*, etc. At the same time, third person singular subjects form the majority of the verb type perception/relational (e.g. *smell*).

Before moving on to the type of stance studied in this work, let us look at one of the many classifications of stance in current linguistic research. The classification proposed by Du Bois (2007: 142—144) identifies four components of stance: evaluation, affect, alignment and epistemicity. Du Bois proposes the following examples:

(1) *That's horrible.* [Evaluation]
(2) *I'm glad.* [Affect]
(3) *I agree.* [Alignment]
(4) *I don't know.* [Epistemicity]

In the two former, the speaker either evaluates or characterizes an object as having a specific value, as in (1), or positions himself/herself affectively, as exemplified in (2). Example (3) displays an instance of alignment, where the speaker positions his/her stance relative to another party. The last example (4) is an instance of epistemic stance and deals with the degree of

commitment speakers or writers engage in the truth of their propositions. Here the speaker expresses weak commitment to his/her proposition by using the epistemic expression *I don't know*, or in other words, applies uncertainty towards what he/she is saying. It is this type of stance that functions as one of the objects of the present study.

Stance-taking has also become a popular object in discourse studies. The interactive features of stance markers in communicative contexts have been closely scrutinized in the works of Hunston and Thompson (2000) and Englebretson (2007). These volumes are specifically dedicated to the discursive and textual functions of evaluative, affective and epistemic realizations of meaning. The present study benefits from these studies in numerous ways. For instance, Hunston (2007: 27—48) emphasizes in her work the importance of combining qualitative and quantitative approaches to stance for a more comprehensive understanding of the phenomenon. In addition, the idea of the intersubjective and interpersonal realization of stance has led Du Bois (2007: 139—182) to create a unified framework for studying interactive stance-taking (elaborated further in Section 2.3.1). Last but not least, Kärkkäinen (2007: 183—219) then implements the idea by investigating the epistemic marker *I guess* as a social act realized through dialogically constructed communication. All these dimensions are also implemented in the present study.

The following section examines epistemic and evidential stance and presents a few studies relevant for the current study.

2.1.1 Previous works on epistemic and evidential verbs

The abundance of epistemic and evidential markers in spoken and written English has been demonstrated in Biber and Finegan (1989), where evidential stance is found to be more frequent than affective markers, which is primarily caused by the more marked grammatical and lexical means of expressing speaker commitment and source of information. The phenomenon covers a wide range of forms, such as epistemic phrases, adverbs, adjectives, verbs, etc. (Kärkkäinen 2003: 20). In this section, I specifically focus on previous studies conducted on stance verbs.

Confusion as to the terminological and conceptual aspects of the two types of modality, epistemic and evidential, has been omnipresent in literature. For instance, Palmer (1986) acknowledges the distinction between commitment and source of evidence, but still gives various

examples of mixed systems. Chafe (1986) treats epistemicity as a sub-category of the larger group of evidentiality, where the latter involves all instances where speakers or writers make references to their attitudes towards knowledge. In her work on the relationship between *seem* and evidentiality, Aijmer (2009) decides to keep the two notions separate, but still acknowledges the overlapping properties of the two. Cornillie (2009), on the other hand, argues against the automatic and seemingly too inclusive or exclusive previous approaches to labeling epistemic and evidential markers, and notes that speaker commitment is rather a result of the interpretation of source than a direct outcome of the mode of knowing. The author also states that the degree of reliability should not be confused with the degree of speaker commitment.

What is relevant for the present study, however, is the acknowledgement of the close and interwoven relationship between the two categories. The present study treats the two constructions as facets of the same phenomenon and argues that just like *it seems* can simultaneously express evidentiality and epistemic necessity (Aijmer 2009), with the help of enough contextual cues, *I think* as a prototypical epistemic marker, can also carry indications of source. Therefore, I believe that epistemicity could hardly be fully understood without acknowledging the source on which knowledge is built.

Epistemic and evidential verbs have primarily been studied in the medium of spoken language (e.g. Biber et al. 1999). Biber et al (1999: 982) find epistemic mental predicates (which they call comment clauses) to be generally rare compared to single adverbs (such as *probably, maybe* and *definitely),* but used with moderate frequency in conversational settings. For instance, in Kärkkäinen's (2003: 37) study of epistemic stance in American English conversations, epistemicity tends to be expressed by a rather limited set of expressions, out of which cognitive and perception verbs form the majority. The three most popular stance markers in her study are *I think*, *he/she said* and *I don't know*.

When linguists investigate epistemic probability and possibility, it has become common practice to position markers indicating these properties on an epistemic scale. Nuyts (2001: 110—111) acknowledges the difficulty in marking the positions of mental predicates on this abstract scale. He nevertheless claims that the epistemic marker *do not know* tends to mark the mid-scalar position, with *know* placed on the 'positive' side indicating certainty, and *doubt* on the 'negative' side of the scale with little or very weak modality. At the same time, he acknowledges the minor role the scale

plays in assigning meaning to the predicates and refers to evidentiality as a concept that eventually determines their semantics. Moreover, it has been suggested that without a thorough look at the context in which single-speaker contributions occur, it is not fully possible to observe the emergent meaning of stance (Martin and White 2005; Kärkkäinen 2006).

The most thoroughly studied mental predicate in the field of epistemic stance is undoubtedly *I think.* The first corpus-based grammar book *Longman Grammar of Spoken and Written English* (1999) lists *I think* as the most frequent predicate construction in both British and American English conversations. Thompson and Mulac (1991) also note that *I think* is the most frequently occurring subject-verb construction alongside *I guess* in their data. What is more, the construction serves as one of the two stance verb phrases in the quantitative part of the present study and is therefore given special attention in the present section. Although well studied, *I think* and mental predicates in general are perceived to possess complex semantic structure (Nuyts 2001: 107), and pose great difficulties in the full understanding of their conceptual and functional abilities in conversational data.

I think in Chafe's (1986: 266) work on evidentiality in English is listed as a verb of belief. Belief is one of the four modes of knowing established by Chafe, where evidence plays a secondary role, or as the author notes, "[…] belief is always based on something other than evidence alone" (Chafe 1986: 266). Capelli (2007: 99), however, claims that particular contextual constraints can accentuate different construals of an epistemic marker, and even such traditionally non-evidential verbs as *I think* and *I doubt* can exhibit indications of source and evidence. While Chafe perceives *I think* to adopt the meanings of belief and opinion (although opinion, he states, is a weaker form of the former), Aijmer (1997) makes a more fine-grained distinction of *I think* and its functional characteristics.

The central topic in Aijmer's (1997: 21—28) investigation of the semantic intent of *I think* is the signaling of either deliberation and authority or tentativeness and uncertainty. Whether the mental verb is deliberative or tentative is determined by its position in the utterance, the presence or absence of the *that*-complementizer and prosodic prominence. Therefore, *I think* is classified as deliberative when it receives prosodic prominence, is positioned initially in the utterance and is followed by the *that*-complementizer. All other cases are classified as tentative.

Kärkkäinen's (2003) study on the interactional functions of epistemic stance in English conversations serves as an answer to previously conducted work on the functions of *I think,* such as Holmes (1990) and Aijmer (1997). The two-way classification of *I think* as either indicating certainty or tentativeness in these studies oversimplifies the various functions the mental predicate possesses in naturally occurring language. Kärkkäinen notes that instead of marking a place on the epistemic continuum between deliberativeness and tentativeness, the cognitive verb rather performs routinized work in interaction and functions as an important discourse organizational marker. For instance, when placed turn-initially, *I think* can have a function of marking the boundaries of a new frame in conversation and indicating the starting point of the current speaker's perspective. Turn-final *I think* often signals turn completion and encourages the interlocutor to provide a response to what was previously presented (see Chapter 5.2.2 for a full classification). Therefore, Kärkkäinen treats *I think* as a fully-fledged discourse marker whose meaning is heavily reliant on its sequential organization and position in larger contextual units.

In Kärkkäinen (2006), the author, who has devoted most of her research on *I think* alone, takes one step forward and argues that stance is the result of situated dialogic interaction. By looking at informal face-to-face conversations, the author finds that stance is a locally constructed phenomenon, where speakers modify their choice of evaluative expressions relative to their conversational partners. Kärkkäinen asserts that stance is a public phenomenon and can only emerge in the course of a communicative activity, which makes it a dynamic concept rather than a static mental state. She finishes with a conclusion that "constructions that are considered prototypically subjective in linguistic theory, such as *I think,* are really intersubjective in nature" (Kärkkäinen 2006: 724). Kärkkäinen, therefore, opens up a new perspective for studying the interactive nature of the predicate, where thorough research needs to be conducted.

Another verb closely investigated in the present study is the evidential verb *seem*. What distinguishes the verb from *I think* is its primary association with evidential information rather than epistemic commitment. The verb is typically referred to as a perception verb and is therefore associated with such predicates as *appear*, *sound*, *look*, etc. Although the verb has traditionally been conceived as solely referring to evidence and the attitudes that speakers and writers have towards it, recent research in the field has also assigned epistemic extensions to the verb (Aikhenvald

2004; Aijmer 2009). As a result, boundaries between the two phenomena have become fuzzier than ever.

Chafe (1986: 266—268) has listed examples with the verb *seem* as representations of evidence acquired through induction and hearsay. The first mode of knowing makes important references to evidence, in which *seem* acts as a marker that typically indicates uncertainty about what is being discussed. Chafe also marks *seem* as a hearsay device, which has been borrowed from induction to indicate knowledge obtained through language. Therefore, *seem* is found to serve different functions in referring to evidence.

What makes the verb particularly interesting is the fact that *seem* can be involved in numerous constructions. Aijmer (2009: 72) suggests seven possible forms:

— *seem* + *that*-clause
It seems that it is raining
— *seem* + *as if (like)*
It seems as if it is raining
— *seem* + *like* noun phrase
He seems like a fool
— *seem* + infinitive
It seems to be raining
— *seem* + adjective (participle)
He seems agitated
— *seem* + *adjective* + *that*-clause
It seems possible that it will be raining
— parenthetic *seem*
It is raining it seems

Additionally, Aijmer notes that since *seem* can also occur with an experiencer as in *it seems to Mary*, the number of possible patterns is much larger.

Due to the diverse variety of syntactic constructions, the perception verb also evokes various meanings. For instance, the different syntactic frames can refer to different types of evidence on which the speaker's knowledge of the world is based. While *it seems that* is found to refer to knowledge acquired through hearsay or verbal transaction of communication, the copular *seem* followed by adjectives and nouns are generally associated with statements based on direct perception (Chafe

1986; Aijmer 2009). What Aijmer emphasizes is the close link between evidence and the degree of certainty assigned to each construction. This means that apart from copular *seem*-constructions that typically make references to appearance, the majority of *seem*-constructions also express probability and certainty. For instance, the construction *it seems to X* is thought to express (subjective) certainty since it displays the speaker's personal opinions or evaluations on the topics being discussed, while *it seems as if* has been found to mean irreality and uncertainty.

The complex nature of epistemic and evidential stance verbs due to their large variety of syntactic constructions and semantic diversity deserves to be studied in more detail. The present study attempts to contribute to the field by approaching stance from two different perspectives (see Section 2.3).

2.2 Studying computer-mediated communication

Since the expansion of the Internet in the late 1980s and early 1990s, computer-mediated communication (CMC) has become a new and appealing source of research. Albeit a relatively new phenomenon, CMC has already generated a great deal of research in psychology, linguistics and communication studies, all of which make use of the medium's impact on language change and the generation of new social roles and virtual communities. Most importantly, studying CMC has opened up a prolific field of research for linguists, providing work with endless opportunities to trace people's use of language in the medium.

CMC is a cover term for the various modes of communication in Internet environments that include such earlier domains as e-mails, chat rooms, bulletin boards, instant messaging, and later additions like blogs, social media sites and voice-over-IP systems. Each mode is characterized by its own set of features due to the technologies on which it is based and the context in which language is generated (Herring 2002: 111—112). The primary parameter for the division of the various modes is synchronicity. Therefore, such modes as instant messaging, where people need to be logged on at the same time, are different from asynchronous modes of CMC, where messages appear and are read with time constraints. For instance, this is characteristic of bulletin board conversations.

The majority of CMC literature has dealt with the kind of language available in the medium, which has typically led to discussions on the medium's classification as either spoken or written. The technology in which CMC occurs categorizes the medium as written[2]. However, it has become clear that a number of the modes carry speech-like features and are therefore closer to informal spoken communication (Lee 2002). Baron (2010: 2) argues that most studies have generalized across different modes of CMC, while patterns of use can vary considerably across these modes. For instance, although during the emergence of electronic mails the mode was thought to represent informality and stylistic disparity, most youngsters today would consider the mode to display a high degree of formality compared to the impersonal nature of instant messaging. A thorough overview of the characteristic features of the various modes of CMC is presented in Crystal (2001). Crystal avoids allocating CMC[3] to any of the two mediums, but rather believes it to be a medium in its own right that successfully combines the properties retrieved from both spoken and written domains to fit its electronic context.

Most linguistic research on CMC up until the present day has been conducted on the special and distinctive features of its language (e.g. Werry 1996, Hård af Segerstad 2002). What has been the focus of CMC is the users' creativity in replacing such paralinguistic cues found in face-to-face interaction as intonation, exclamations and emphasis with features relevant for the medium. Also, the emergence of such abbreviations as *ttyl* ('talk to you later') and *lol* ('laughed out loud') are properties that have emerged from CMC and are popularly used in the English language. It is above all because of these special features of CMC that the phenomenon is thought of as a leading source behind language change and development. Again, linguists working with CMC have to consider the diverse 'dialects' present in all of its modes and approach the domains accordingly. For instance, Lewin and Donner (2002) have found in their work on bulletin boards that features that most often creep up in forum conversations are special CMC features, such as non-standard spelling (e.g. *yep*, *nope*) and emphasis (typically through capitalization). What the authors, however, admit is the necessity for a sociolinguistic approach to the modes to uncover the demographic differences assumed to play a role in CMC.

---

[2] Except for voice-over-IP systems, such as Skype.

[3] Crystal refers to CMC as Netspeak, which he prefers over the various names given to the medium (2001: 19).

A growing body of research has become more inclined towards studying the sociolinguistic aspects of CMC, with the aim to highlight the existence of social and linguistic diversity in various modes of the phenomenon (Androutsopoulos 2006). A transformation from 'language of CMC' to computer-mediated discourse as presented by Herring (2004) is a strong indication of the shift in focus. One of the main objects of study has become the existence of online or virtual communities, which are thought to be the central aspect in explaining the social behavior of Internet users. For instance, Herring (2004) has established six dimensions, which she believes distinguish authentic virtual communities from the rest. Among others, these include the formation of social roles, rituals and hierarchies, a notion that carries strong importance in the operationalization of social hierarchy in the present study (see Chapter 3.1.3). From the viewpoint of virtual communities, discourse in CMC can be studied relative to a number of paradigms, such as online ethnography, language variation, social interaction and identity, and multilingualism (Androutsopoulos 2006). The present study is concerned with social interaction between members from different hierarchical ranks.

2.2.1 Power and social hierarchy in online communities

A number of studies on CMC have outlined the existence of social roles and power in virtual communities, each of which plays an important role in the intricate web of social responsibilities (e.g. Spears and Lea 1994; Herring 2003). These roles that Internet users find themselves in are believed to have essential effect on the language they use.

CMC has been found to have strong influence on the distinct social behavior of Internet users (Postmes et al. 2001; Herring 2002). The first and most obvious reason for this lies in their relative anonymity. It is mostly the absence of such social cues as age, gender, race, etc. that differentiates the medium from face-to-face communication. This might cause the occurrence of such boorish behavior as flaming[4] and hostility towards fellow users. However, in modes where anonymity is low and users have established a community where ties have been formed, such egalitarian aggression is less problematic. Herring (2002: 137—138) notes

[4] "Flames are lengthy aggressive utterances, related to a specific topic and directed at an individual recipient" (Crystal 2001: 58)

that this mostly happens in asynchronous modes of CMC such as bulletin boards, where a majority of discussions are led by a small number of users with a majority of users simply acting as readers. Herring adds that frequent participation is therefore not solely a matter of anonymity but self-confidence and social courage.

However, in such asynchronous modes of CMC as bulletin boards, hierarchy is often built in the domain, which acts as a strong trigger for the emergence of social roles. Such domains often empower their users by institutionalizing their position in the community (Kolko and Reid 1998). This is exemplified by the existence of such privileged roles as moderators on bulletin boards, Wizards on MUDs[5] and operators on IRC (Internet Relay Chat). As a result, the users' socioeconomic information is often overpowered by the roles established within the boundaries of the online community. The main factor that eventually starts playing a key role is experience, in which case experienced members are opposed to "newbies", who tend to receive less respectful treatment from the former (Herring 2002: 138).

In bulletin boards, the privileged role is often enjoyed by moderators. Moderators are the core of the board and employ a managerial role in the community. They also have a considerable amount of power. For instance, they have the possibility of editing and deleting messages, cut out flaming, spam, illegal advertisement and other undesirable material. What is more, in Marcoccia's (2003) framework of participation roles in Internet newsgroups, the author has identified various characteristic features of participants who conduct the conversation group (and whom he calls 'hosts'). One of these include the tendency for hosts to threaten other senders' faces and display authority towards users who do not follow the board's rules. Reid (1999) establishes in her study on hierarchy and power on MUDs that despite the apparent freedom in cyberspace, such modes display a great degree of social control and power. The author also adds that the desire to achieve Wizardship (the privileged position on MUDs) encourages members to invest time and selfless dedication; all this to reach the highest level of the hierarchy and achieve recognition and praise from the rest of the community.

The present study attempts to contribute to the field of power and CMC and looks at the expression of stance as a device for highlighting hierarchies between users of an online bulletin board.

---

[5] MUDs are "networked, multi-user virtual reality systems [...]" (Reid 1999: 107)

2.3 Theoretical assumptions behind the methodologies

2.3.1 Dialogicality

In dialogicality, the concept of an 'ideal dialogue' that is free from any interactional predicaments, such as discontinuities, ambiguities, vagueness and miscommunication, is considered to be a myth. In addition, with Cognitive and Functional linguistics, language is no longer viewed as an autonomous system of words and symbols but a result of human cognition and interaction with our experience and existence in the world. Therefore, it is argued that the idea of an 'ideal dialogue' cannot serve as the basis for any empirical research due to the dialogical constitution of interaction (Linell 2009: 5). The belief has its roots in the works of Bakhtin (1981), whose ideas emerge from the perspective that Self is constituted through Other and every contribution made by the former is dialogically constructed with the latter. This means that the Self-Other interdependency in discourse is the central issue in how interaction unfolds in communicative contexts. These ideas form the basis of the qualitative part of the present study, in which stance is investigated using dialogical discourse analysis. It is a method that retrieves its basic assumptions from the studies of dialogicality, which are elaborated further in the following paragraphs.

Dialogicality as studied by such scholars as Marková (2003), Linell (2009), Du Bois (2007) and others, assumes that communication does not manifest itself in individual contributions produced by autonomous speakers and writers, but rather forms a complex maze of intersubjective positions adopted by a number of speakers and writers (Du Bois 2007; Marková et al. 2007). Conversations are therefore seen as dynamic processes in which meanings are constructed as conversational participants advance in their communicative acts, which results in the rejection of fixed and stable thoughts, ideas and positions. In addition to the alignment of contributions relative to previously occurring turn units, dialogicality is also concerned with speakers' awareness of what might follow. This motivates speakers and writers to direct their utterances at possible future replies. With others in mind, speakers and writers are therefore interdependent with the cognition and language of their conversational partners, which highlights the role of humans as social beings (Linell 2009: 148).

Dialogicality is closely associated with sociolinguistics and the socio-cultural space of sense-making (Marková et al. 2007). Dialogical researchers believe that people's ideas, opinions and perspectives are at least partly guided by the social perspectives dominant in the community to which they belong. Marková et al. (2007) have termed it 'socially shared knowledge'. Socially shared knowledge is perhaps best understood in terms of the triangular relation Ego-Alter-Object, in which the latter as a topic of discussion is socially and inclusively generated in the communicative act between Ego and Alter or Self and Other (Marková et al. 2007: 22). The framing of the Object between the conversational participants rarely displays neutral and homogeneous representations. People as social beings are representative of a diverse and heterogeneous array of positions and attitudes towards the world, which often triggers clashes and confrontations of ideas, opinions and beliefs.

Table 1. Stance diagraph indicating the alignment of stance between two interlocutors (Du Bois 2007: 166)

#	**Speaker**	**Stance Subject**	**Positions/ Evaluates**	**Stance Object**	**Aligns**
1	SAM;	I1	don't like	those	
3	ANGELA;	I2	don't{like}	{those}	either

Evaluation and stance are central in the study of dialogicality, since they are believed to be shaped by the intricate interplay of conversational collaborations by dialogical co-participants (Du Bois 2007: 141). However, Du Bois (2007: 174) notes the lack of a unified framework for expressing stance and proposes "a minimum structure of stance as dialogic action," or what he calls the 'stance triangle'. The model is based on the dialogical prominence in the construction of stance, where speakers explicitly imitate the syntax of previous parties. The three components Du Bois proposes for the interpretation of stance are evaluation, positioning and alignment. In other words, when expressing stance, the speaker evaluates an object, positions the subject (which often refers to the speaker himself/herself), and finally aligns with other stance-takers (Table 1, above). The table represents a stance diagraph, where the second speaker Angela's stance clearly builds on the previously expressed stance by Sam, which is realized by the lexical item *either*.

Another important objective of dialogicality that closely corresponds to the goals of the present study is the negotiation of power and social

hierarchy. In fact, power and powerlessness are often discussed in the literature of dialogicality (e.g. Linell 2009). Linell approaches power and powerlessness the same way he approaches communication in general by stating that power is also the result of interaction: "Power emerges from interaction, and is executed in and through interaction […]" (Linell 2009: 216). Linell also adds that both power and powerlessness are conceptually intertwined, in which case the interlocutor with power cannot exercise it unless the other party succumbs to domination and silence.

What distinguishes dialogicality from other approaches to text and talk is its strong interest in the cognitive and conceptual nature of communication. This differentiates it from such models as Conversation Analysis (Sacks, Schegloff, and Jefferson 1974) that solely studies external aspects of discourse. Dialogicality therefore makes assumptions about the strongly intertwined relationship between communication and cognition (Linell 2009: 15). Dialogically constructed interaction has been studied in a variety of contexts, such as focus group discussions on social issues (Marková et al. 2007), misunderstanding and miscommunication in immigration divisions of police and health institutions (Linell 1995), etc.

It should be noted, however, that this study does not attempt to give a full account of the complex nature of Ego and Alter and their interdependence and importance in communication. It is merely concerned with the main perspectives of dialogicality, in which communication is seen as a complex maze of intersubjective contributions where stance is locally constructed by two or more co-participants.

2.3.2 Quantitative Cognitive Semantics

The following section deals with quantitative corpus-driven Cognitive Semantics and most specifically with multifactorial usage-feature analysis. Multifactorial usage-feature analysis is a method used for the investigation of epistemic stance in the quantitative part of the study, and therefore, I give an overview of the basic assumptions behind the approach.

The emergence of Cognitive Linguistics in the 1980s brought about major changes in the description of language. What had so far been explained through structural properties inherent in the symbolic system of language, now started to be defined by such features as cognitive processes, the principle of categorization, interactional devices, etc. Instead of syntax, the central object of study became semantics and the mapping between

meaning and grammatical structures. Although the first wave of cognitive linguists (e.g. Lakoff 1987; Langacker 1987) produced an inherently empirical linguistic framework, these implications were never tested with empirical research. However, the second generation of cognitive linguists saw the implications of the framework and the 1990s prompted a swift emergence of empirical research in Cognitive Linguistics (e.g. Schmid 1993; Geeraerts, Grondelaers, and Bakema 1994). This is a crucial landmark in quantitative Cognitive Semantics, where a large number of examples are studied using the multivariate modeling of semantic meaning.

The central idea of quantitative corpus-driven semantic research is the belief that language acquisition and change are governed by language use, i.e. it is usage-based. The usage-based framework assumes that structure is the result of frequency and repetition of syntactic and semantic constructions in language, and most importantly, in natural discourse. In Langacker's (1987: 48) terms, some lexical items are more readily recognizable than others due to their larger degree of *entrenchment*, "which reflects the frequency of […] previous activation and determines the likelihood of […] subsequent activation." The linguistic units with frequent occurrences become more strongly entrenched in our memory, and as a result, determine the characteristic constructions used in speech communities at large. Substantial databases or corpora of naturally occurring language have enabled researchers to measure the frequency of prefabricated chunks of lexical units, which have been structuralized in speech and writing. They have also favored the emergence of exploratory studies to determine the current patterns of linguistic expressions.

With entrenchment, Langacker proposes a quantifiable definition of grammaticality, in which the individual's frequency of use of a pattern is extended to the whole speech community (Glynn 2010a: 6). As a result, Langacker has found a way to operationalize grammaticality. Linguists working in quantitative Cognitive Semantics are similarly looking for ways to operationalize an inherently intersubjective object of study, namely semantics. There are two main schools working towards the goal, which include the multifactorial school (Divjak 2006; Glynn 2010b; Speelman and Geeraerts 2010) and the collocation-based school (Stefanowitsch and Gries 2003; Wulff, Stefanowitsch, and Gries 2007). The latter deals with the degree of repulsion and attraction between words, while the former displays the multidimensionality of linguistic structuring. In this paper, I focus on the multifactorial approach to language.

Multifactorial analyses first of all reflect the multidimensionality of and interaction between morphology, syntax, lexis and prosody, and secondly, help to quantify the degree to which various factors contribute to the meaning of grammatical and lexical constructions (Geeraerts, Grondelaers, and Bakema 1994; Divjak 2006; Glynn 2010b). Perhaps the most important advantage of the analysis is its suitability for applying statistical analysis to a large number of examples, which significantly improves the speed and quality of multifactorial modeling (Gries 2003; Glynn 2009; Speelman and Geeraerts 2010). With the adoption of both exploratory and confirmatory statistical tools, researchers are then able to first visualize the patterns in their data and secondly test the statistical significance and predictive power of their sample. By doing so, linguistics and especially the cognitive framework is tentatively taking a step towards establishing itself as an empirical approach to science-making, where hypotheses can be supported or falsified by using numerical measures.

In the field of epistemic and evidential verbs and the Cognitive Linguistics framework, quantitative approaches have been scarce. One of the exceptions is Krawczak and Glynn (2011), whose investigation of social cognition and epistemic parentheticals sheds light on the socio-cognitive realization of epistemic stance in Internet weblogs. The present study attempts to contribute to the field by combining the two approaches, dialogical discourse analysis and multifactorial usage-feature analysis, to uncover the impact of social status on language use.

Chapter 3
The data

This chapter describes the data retrieved from the online bulletin board rsg.net. First, I highlight the thematic and structural constitution of rsg.net (Section 3.1.2). I then show that rsg.net and bulletin boards in general bear in their scope an important attribute, which makes it possible to operationalize an inherently intersubjective phenomenon, social rank (3.1.3). Additionally, sections 3.2 and 3.3 introduce the processes of data collection for qualitative and quantitative analyses respectively. The data for the qualitative study contains two threads of 8,678 words where epistemic and evidential verbs are studied in their interactive context. The quantitative part of the study retrieves its data from the user profiles of 12 native speakers of Australian and British English to identify the discursive patterns of authoritative and tentative stance. Although the two methods adopted in the present study, dialogical discourse analysis and multifactorial usage-feature analysis, are based on data from the same bulletin board, different techniques have to be adopted to adapt to the constraints and possibilities posed by each approach.

3.1 Computer-mediated communication

3.1.1 Bulletin boards

The Oxford English Dictionary (OED) gives *bulletin board* the following definition: “a computer-based system giving users access from remote terminals to text and programs contributed by one another and stored centrally” (“bulletin board”). The earliest instance of the word that refers to electronically stored discussion sites as opposed to concrete bulletins

displaying public news and events originates from 1979. Although broad, OED's definition of bulletin boards captures two centrally important features of this study: (1) bulletin boards are computer-based and form part of the large network of CMC, and (2) users of the board make textual contributions to centrally stored sites from distant parts of the world. What the definition fails to capture, however, is the essential function of such virtual forums: the exchange of knowledge and views on the topics being discussed.

Internet bulletin boards are asynchronous modes of CMC, meaning that texts entered to the system occur with time constraints and members of the board do not have to be online at the same time. The main difference between bulletin boards and other modes of CMC lies in their ready accessibility for all Internet users, or as Claridge states, "forums […] are part of the public world-wide-web space, look like 'normal' web sites and can be visited and read by any internet user at any time […]" (2007: 88). Undoubtedly, its easy accessibility is an essential reason for the heteroglossic and highly diverse language found in such virtual communities.

A crucial feature of bulletin boards that has strongly tempted researchers working on CMC to treat the mode as an oral rather than a written representation of communicative exchange is the abundant existence of informal language. Claridge (2007: 89) claims that due to the significant role that information exchange, evaluation and interaction play in forum communication, the investigation of interactive stance markers seems to be an efficient first approach to bulletin boards. In addition, Davis and Brewer (1997: 153) have noted in their search for linguistic features characteristic to this particular mode of CMC an overwhelming use of the personal pronoun *I*, the adoption of *it* as in *it seems to me*, and the implementation of private verbs such as *think*, *feel* and *know*. Therefore, previous studies on the nature of bulletin board conversations unequivocally confirm the high frequency of epistemic and evidential markers in the mode, and most importantly the necessity for more research in the field.

3.1.2 Rsg.net

Rsg.net is a recreational Internet bulletin board that is mostly used by young female gymnastics fans, although the proportion of male fans has

been on the rise in recent years. It is specialized in rhythmic gymnastics, a sport where individual or group gymnasts perform to a piece of music while manipulating one of five apparatuses: rope, hoop, ball, clubs or ribbon. As of 13 March 2013, the total number of users of the board is 4,790. They have contributed over 300,000 entries since the board was opened in 2003. Therefore, the board is a fertile source of human interaction in computer-mediated environments that has not yet been investigated.

Table 2. Rsg.net as of 15 March 2013, 11:00 A.M.

	Number of threads	**Number of posts**
News & Announcements	11	11
The Gym	1,196	18,607
Code of Points	579	8,195
Events & Results	1,572	73,360
Gymnasts & Groups	639	18,206
Leotards & Equipment	1,095	14,099
Music & Editing	486	24,059
Gymnastic Photos	634	40,472
Photo Announcements	5	58
Gymnastic Videos	1,092	48,161
Video Announcements	7	110
Old Photos & Videos	3,048	34,614
Off Topic	548	29,835
RSG.net	57	1,202
Total	**10,969 threads**	**310,989 posts**

Participation on the board has proved to be individually beneficial, allowing for self-expression and formation of relationships with other fans. Largely because the forum unites people with shared interests, its most loyal users have formed a relatively tightly knit Internet community, where knowledge and ideas are generously exchanged. Nonetheless, despite sharing a passion for rhythmic gymnastics, criticism and conflicts have almost become an everyday matter on the board.

The structure and organization of bulletin boards follow a rather firmly pre-established layout. With the expansion and growth in popularity of Internet domains, the traditional forum-thread-message outline has been replaced with a more complex categorization of themes. Rsg.net is no exception in that area, and due to the large number of possible fields of discussion in gymnastics and better navigation between them, the posts in

rsg.net have been divided into 14 groups: News & Announcements, The Gym, Code of Points, Events & Results, Gymnasts & Groups, Leotards & Equipment, Music & Editing, Gymnastic Photos, Photo Announcements, Gymnastic Videos, Video Announcements, Old Photos & Videos, Off Topic and RSG.net. After entering one of the topics, registered users are able to start their own threads within the borders of the given theme or contribute to the existing ones initiated by others. The most popular topic among rsg.net users in terms of the number of posts is Events & Results (see Table 2, above), where the latest news of various competitions and performances are meticulously scrutinized. Other popular topics include photo and video sections and Off Topic, since not everything is about rhythmic gymnastics.

3.1.3 The operationalization of social hierarchy

The operationalization of social hierarchy is a crucial step in determining the social dimension of stance-taking in the present study. As mentioned above, stance-taking is both a linguistic and a social act. Giving insights into its variation in a mode, where members are *a priori* considered to be equal, might shed a novel light on the socio-cognitive realization of stance. Since language serves as one of the few (if not the only) indicators of users' personae in online anonymous communication, "[it] becomes the primary means of establishing and maintaining group membership and identity" (Crystal 2001: 156). Therefore, in the present study it is believed that language variation is primarily influenced by the social roles the members of the community are assigned to play.

In Herring's (2004) framework of CMC, the author poses six criteria for Internet groups to be categorized as authentic virtual communities. One of them is the formation of social roles. However, due to the incomplete and often unreliable socioeconomic background of Internet users, such as age, gender, race, etc., researchers need to turn to other ways of operationalizing social status.

In this study, the impact of power and social hierarchy on language variation is determined on the basis of two criteria inherent in the structure of rsg.net and bulletin boards in general: (1) the assignment of privileged status based on the members' contribution to the domain, and (2) their activity and involvement in rsg.net discussions. As a result, social hierarchy is operationalized by dividing members of rsg.net into three ranks

according to their status (Rank 1) and activity (Ranks 2 & 3) on the board (Table 3). The first criterion solely applies to the Rank 1 users or *moderators*, who are categorized as such irrespective of their activity on the board. Members from Ranks 2 and 3, however, are identified based on their activity or the number of messages contributed to the board. The two ranks are distinguished by a landmark of 700 messages, which means that rsg.net members with more than 700 messages are categorized as *hosts* and members with less than 700 messages as *casual senders*. The number is chosen after a thorough observation of the relative activity on the board.

Table 3. The operationalization of social status on rsg.net

Rank	Title	Description
1.	Moderators	users who have been promoted to keep order on the board, number of posts not taken into account
2.	Hosts	users whose contribution exceeds 700 posts
3.	Casual senders	users who have contributed fewer that 700 posts since joining the board

The division is partly retrieved from Marcoccia (2004), whose distinction between simple readers, casual senders and hosts is modified to fit the organization of the present study. For example, Marcoccia's simple readers, who only act as readers of the posts, do not leave behind any written traces that could be linguistically investigated, and are therefore automatically excluded from the categorization. It is replaced (or brought down one level) by the category of *casual senders*. The rank above casual senders will be named *hosts* that represents a relatively modest proportion of active and experienced members, who nevertheless do not enjoy the privileges and power of *moderators*. In Marcoccia's categorization, *hosts* are distinguished from other users by a selection of interactional behaviors, such as an active participation in sending and answering to messages, conducting conversations, playing the role of experts inside the community, and being on friendly terms with other members of the board. Surely, one can argue that these features could easily be assigned to the highest level of the current classification. However, in rsg.net, *moderators* do not only make frequent contributions to the board, but are also given a privileged role by the administrator of the board to keep order in the community by filtering offensive and obscene messages and spam, move misplaced posts to relevant thematic groups, and ban misbehaving users. Therefore, the last and most powerful rank consists of members whose trustworthiness as well

as active participation on the board has given them considerable power over the rest of the community.

The reason for choosing the number of posts as an indication of activity lies in the members' possession and distribution of valuable information, news and gossip about the stars of the sport, and behind-the-scenes reports from local and international competitions – all of which are believed to be the reason behind the board's world-wide popularity. The assumption that it is above all information that sets some members higher on the hierarchical scale than others is also in accordance with Marcoccia's interactional behaviors presented above. In addition, according to Herring (2002: 137—138), the majority of discussions evolved in asynchronous Internet modes are dominated by a small number of users. These members are associated with such characteristics as self-confidence and perceived entitlement and are often opposed to less active and novel users, or 'newbies', who enjoy fewer rights by sometimes even receiving less respectful treatment from the core members of the board. The present study, therefore, predicts that activity and the exchange of information is the primary factor in differentiating *hosts* from *casual senders*.

In rsg.net, the operationalization of social hierarchy is determined by status and the overall number of messages posted on the board. This information can easily be retrieved from each user's profile. Unlike such socioeconomic information as location, occupation and interests, rsg.net automatically saves the date of registration, total posts provided by the user, and the average number and percentage of posts per day. Figure 1 gives an overview of how this information can be drawn from each member's profile. This fan with a username *Tahnee* has contributed a total of 3,272 posts since 11 January 2004, making him/her a relatively experienced member of the board, which is also determined by his/her status as a moderator (indicated on the top-left corner).

Figure 1. Rsg.net user profile of *Tahnee*

3.2. Data collection: qualitative study

The present study is divided into two methodological frameworks. The first part of the study adopts a qualitative approach to corpus linguistics, where two controversial discussions or threads are studied to give an in-depth overview of epistemic and evidential stance constructions in discourse. The methodology, dialogical discourse analysis, retrieves its basic assumptions from the studies of dialogicality (Linell 1998; Marková 2003), in which stance-taking is seen as a usage-based communicative act that is mutually constructed between two or more co-participants, as well as a lexical property through which these participants constitute social relationships between one another.

In the next section, I discuss the methodological considerations involved in the selection of two discussion threads and present a step-by-step description of their retrieval from rsg.net. The threads are consequently analyzed in Chapter 4 using a dialogical discourse analysis.

3.2.1 The retrieval of two threads from rsg.net

The discussion threads chosen for the study need to display a high degree of controversy and conflict and exhibit a great variety of various techniques of stance-taking. After some investigation of a number of threads, two discussions were retrieved from the domain: (1) *Why does Merkulova*

receive so many critics? and (2) *Can I still become an Olympian at age 19?* The two topics display thematic and topical variation as well as differences in the people being talked about. For a more comprehensive account on the retrieval of the two threads, both discussions are described separately. However, it should be noted that in the qualitative study, regional variation has not been controlled for due to (1) the difficulties in extracting rsg.net conversations with members from the same language group, and (2) our interest in studying the dynamic behavior of epistemic and evidential stance verbs relative to conversational co-participants as opposed to the construction of stance *per se*. Nevertheless, the language-specific construction of epistemic and evidential stance is taken into account in the quantitative part of the study (see Chapter 3.3 for the determination of regional dialects).

Thread 1. Why does Merkulova receive so many critics?

The first thread focuses on a controversial young Russian gymnast, namely Aleksandra Merkulova, whose high marks in international competitions are believed to be the result of bribing, or over-scoring, which is the term often used by members of rsg.net. Although bulletin boards are easily accessible for all Internet users, Aleksandra Merkulova is not a registered member of the board and there are no records of her visits to rsg.net. Therefore, the protagonist of the thread acts as an invisible force behind the arguments and conflicts unfolding in the course of the discussion, and it is assumed that the messages written by members of rsg.net are not produced with much concern about Merkulova's face wants, despite the fact that concerns are being expressed.

The number of posts in this thread is 151, which contain 8,678 words. The discussion extends over two and a half months, from 6 June 2012 to 24 August 2012. The entries and their metalinguistic information were extracted on 6 November 2012. The metalinguistic information includes the name of the sender, the user's personal information (if provided) and the date of posting. Since analyzing the discursive behavior of all verbs of epistemic and evidential commitment would go beyond the scope of the present paper, two extracts of 845 words (572 and 273 respectively) were obtained from the thread (see Extracts 1 and 2 in the Appendix). Due to the fragmentary nature of bulletin board discussions, in which case one thread can display the emergence of multiple conversations or users can be involved in various conversations (Marcoccia 2004: 120), the messages were selected relative to their topical continuation and homogeneity. Other

criteria include a sufficient representativeness of epistemic and evidential predicates and the presence of all three social ranks.

Table 4. First post in thread *Why does Merkulova receive so many critics?*

Author	**Message**
Brivido	**Posted:** Wed, 6-Jun-2012 19:38 **Post subject:** Why does Merkulova receive so many critics?
Joined: 19 Apr 2012	When she was a junior, everybody loved her. Now, everybody says she's a clown. Can you explain from where comes all this hatred?
Posts: 43	

Extract (1) of the first thread displays the initial post with entries clearly directed at its thematic content. The number of conversational participants in the thread is five and the ranks represented are 2 and 3. Collectively, they have produced six posts. Extract (2) represents a subtopic generated in course of the overarching theme. The number of posts here is four, corresponding to the number of members involved in the discussion. As opposed to the previous extract, the present excerpt represents rsg.net members from all three ranks. Altogether, the two extracts constitute the discussion unfolding in the thread *Why does Merkulova receive so many critics?*

The thread was retrieved from the rsg.net group called Gymnasts & Groups that carries the following tagline: "What would this sport be without our heroes: the gymnasts!" Threads in this group typically revolve around a particular gymnast or group and contain news, reports and opinions about 'the heroes' of gymnastics. It should be borne in mind that the initial title of the thread was *Why does everyone hate Merkulova*, but the users' continuous complaints about the use of such a powerful word as *hate* ("Hate is a strong word") and the ensuing effect the title might have on the particular gymnast ("[…] for respect to Aleksandra, please change the title") forced moderators to change the title to something more neutral. This suggests that although there is little chance for the protagonist to ever read the thread, the sensitivity of the issue is being acknowledged among rsg.net members themselves. Nevertheless, the word *hatred* still remains in the first post of the discussion, posted by the Rank 3 user *Brivido* (Table 4, above).

Thread 2. Can I still become an Olympian at age 19?

The second thread revolves around an inquiry posed by the Rank 3 user *OlympianRG* and deals with his/her chances of participating in the Olympic Games, when starting rhythmic gymnastics trainings at the age of 19. Today, girls can start practicing gymnastics as early as 3—4 years old and typically enjoy success shortly after entering the 'senior' category at 15, which can prompt a retirement before one reaches 20. In fact, gymnasts competing on a professional level in their twenties have become a rare phenomenon in a sport that can have deteriorating effects on their health and body. Therefore, for the majority of rsg.net members active in rhythmic gymnastics, the idea of starting at the age of 19 and dreaming of being part of what can be considered the greatest sports event in the world is an unquestionable trigger for generating rather offensive judgments.

The thread consists of 36 posts with 5,168 words. The 'life span' of the topic ranges from 13 March 2011 to 29 December 2012 and the posts were extracted on 29 May 2013. From the thread, one extract of 1,338 words is selected for a thorough investigation of stance (see Extract 3 in the Appendix). The extract is thematically homogeneous, i.e., it includes the first post with a number of entries addressed directly to its content. The number of members who take part in the discussion is six, who represent all three hierarchically different ranks. As each member has produced one message, the number of posts in the thread corresponds to the number of participants.

This thread was extracted from a group named The Gym, which is characterized by the following tagline: "What's up in Rhythmic Gymnastic? News, rumours, fans – you'll find them here!" In short, the group contains a large variety of possible discussions, ranging from tips for improving flexibility to predictions for upcoming competitions. This particular thread is posted by the Rank 3 user *OlympianRG* (Table 5, below), who unlike Aleksandra Merkulova, is not an acknowledged and well-known gymnast, but a casual member of rsg.net asking for advice from his/her associates.

Altogether 16 posts with 2,183 words are thoroughly studied in the qualitative part of the present work (see Chapter 4 for analysis). By investigating the two threads, I hope to find enough stance markers to answer the question of how rsg.net members from three hierarchically different ranks mitigate and reinforce the strength of their propositions relative to their conversational co-participants.

Table 5. First post in thread *Can I still become an Olympian at age 19?*

Author	Message
OlympianRG **Joined:** 13 Mar 2011 **Posts:** 1	**Posted:** Sun, 13-Mar-2011 19:05 **Post subject:** Can I still become an Olympian at age 19? Ok, I know it sounds old to start RG at age 19, but I am really determined and prepared to work hard. Over the past few years I've become seriously enamored with this sport. I think it's one of the most beautiful sports there is. Please don't be confused by the 'Oympian' in the title, I know aiming for gold might not be possible, but just getting in the Olympics would already be heaven for me. I'm a perfectionist at everything that I do. I haven't followed any RG courses yet, as I wanted to focus on school first and my parents wouldn't let me. They said I still 'd have plenty of time after high school. I can already do all the splits with a slight oversplit on the left. I weigh 49kg and am 1.75 meter, which is the same as Anna Bessonova . I can also do a bridge and get my hands to about 15cm of my feet. I got a bit scared when I see people retiring at 23 years old . Why do they do that? Can't they still get better at RG? I would never give up. I'm from Belgium, so I'm looking for the best place to practice and a good coach. Could you guys please help me. I know it might take years and years to get to my goal, but I'm not a quitter and am prepared to work very hard. I'm already stretching about 5 hours a day.

3.3 Data collection: quantitative study

In the following section, I describe the process of data collection for the quantitative part of the study. The procedure differs significantly from the one described in 3.2. The reason for this lies in the adoption of a quantitative as opposed to a qualitative approach to epistemic and evidential stance and the analysis of only two verbs, *think* and *seem,* in their immediate context. The verbs are chosen for their (1) frequent occurrence in rsg.net messages based on prior observations, and (2) their tendency to refer to both epistemic commitment and source of evidence. The verbs are extracted from the user profiles of 12 native speakers of Australian and British English. The analysis adopted for all instances of the two verbs is multifactorial usage-feature analysis (Geeraerts, Grondelaers, and Bakema 1994; Gries 2003; Divjak 2006; Glynn 2009). It is a relatively

recent development in Cognitive Semantics, which helps linguists to quantify the degree to which various formal, semantic and extra-linguistic factors contribute to the meaning of grammatical and lexical items.

3.3.1 The retrieval of think *and* seem *from rsg.net*

The collection of data for a successful analysis of the epistemic *think* and evidential *seem* using a quantitative corpus-driven approach to Cognitive Semantics consists of three main steps:

1. The determination of regional dialects (Australian and British English)
2. The extraction of posts from 12 user profiles
3. The identification of utterances containing the epistemic *think* and evidential *seem*

In the following sections, I give an overview of all three steps in more detail.

Step 1. The determination of native speakers of Australian and British English

Since expression of stance is a language-specific phenomenon, controlling for regional variation is the first concern in the process of data collection. As mentioned above, rsg.net unites people from all over the world with different backgrounds, and most importantly, from distinct language groups. The majority of members are non-native speakers of English, typically from European countries, where rhythmic gymnastics has the highest number of practitioners as well as longer traditions in competitive gymnastics than anywhere else in the world. It is therefore important to control for regional variation to minimize the affect of cultural variation on stance-taking, which is believed to have considerable influence on the phenomenon.

As the main goal of the study is to identify differences in the language of rsg.net members from three hierarchically different ranks, it is important to control for an equal number of members and messages in all three ranks. As the number of moderators or Rank 1 users is limited to a mere 12 members (including the administrator of the site), it is the availability of Rank 1 users with a shared native language that ultimately determines the

dialect. As altogether three users with British and Australian English as their first language could be identified from the list, the two regional dialects were selected.

Earlier in the chapter, I discussed the lack of relevant socioeconomic information available for researchers working on CMC. How is dialect identified with a limited access to such information in rsg.net? The identification of Australian and British English in the present study is determined using three cues:

- Personal information about the country of origin provided on user profiles
- Contextual cues in messages, e.g. *In Australia we…*
- The information known to the author of the paper as a long-term user of the same bulletin board

As the insertion of personal information on their profile is an option for rsg.net users, relying on user profiles for an indication of the country of origin is not sufficient. Therefore, an additional examination of messages posted by rsg.net members is a necessary step. For this reason, such topics as *Popular sport in your country?* and other inquiries about traditions and customs in the countries represented in rsg.net were preferred. These topics typically tend to give relevant information about the users' country of origin. Last but not least, as an active user of rsg.net for nearly ten years, the author of the present study has established close relationships with a number of members of the board and therefore possesses valuable socioeconomic information beneficial for the present study. Rsg.net members often meet in real-life settings, such as international competitions, gala shows and training camps, and therefore the information obtained from these encounters is considered to be highly reliable.

The importance of personal contacts and the examination of messages are exemplified in the case of the Rank 2 user *Jonathan,* whose user profile states France to be his country of origin. However, since the user is an active and rather verbal member of rsg.net as well as a well-known gymnastics fan beyond the boundaries of the virtual community, *Jonathan* is known to be a British citizen, who currently resides in Paris. In addition to an abundance of England-related messages and such contextual cues as demonstrated in example (1), I can confidently say that *Jonathan* is a native speaker of British English.

(1) *It's the general public that votes, neighbours vote for each other, so it's nigh-on impossible for the UK to win… seeing as though* ***we're*** *an Island*

Step 2. The extraction of posts from 12 user profiles

Due to the limited number of Rank 1 users from the same language group, only three members from the highest rank can be selected. Similarly, three users are chosen from Rank 2 as the number of posts roughly corresponds to the number retrieved from Rank 1. However, Rank 3 users are represented with six members due to their limited number of contributions to the board. It is acknowledged that the small representativeness of members from each rank is a weakness of the present study and might create results that are affected by individual variation. With the adoption of a model with mixed effects, it is possible to account for individual variation in future research.

The maximum number of posts extracted from one user profile is set to 750, which is expected to give a sufficient number of instances of *think* and *seem* for subsequent quantitative analyses. If the user has contributed less than 750 posts, as is the case with one user from Rank 1 and all users from Rank 3, the entire set is extracted from the profiles. The time period during which the messages were produced ranges from 27 October 2003 until 6 November 2012. Fortunately, the three moderators chosen for the study have been holding the position for a relatively long time and therefore all 750 messages are posted after gaining the privileged position. For better understanding, Table 6 (below) gives an overview of 12 users from three hierarchically different ranks with their rsg.net usernames, dialects and the number of posts retrieved from the board.

As already mentioned, Rank 3 users are classified as such if their total number of posts does not exceed 700. As Table 6 indicates, all users have contributed significantly less than 700 messages and can therefore be confidently treated as novel users of the board. When it comes to Ranks 1 and 2, a numerical discrepancy in the last column becomes strikingly apparent. There is one user from both ranks (*singlo* and *RGribbonqueen*), whose number of posts is clearly lower compared to the rest of the members from corresponding ranks. However, as summarized in Table 3 above, moderators are classified as such only in terms of their status on the board, irrespective of the number of messages contributed to the domain. Therefore, *singlo* is still considered to be a rightful member of his/her rank.

Table 6. Summary of posts retrieved from 12 user profiles

Rank	Username	Dialect[6]	Number of posts
RANK 1	*Tahnee*	AusEng	750
	Storm	AusEng	750
	singlo	BrEng	145
			1645 posts
RANK 2	*Jonathan*	BrEng	750
	AnnaBessonovaNumber1Fan	AusEng	750
	RGribbonqueen	BrEng	242
			1742 posts
RANK 3	*rg_chik*	AusEng	408
	Bri	AusEng	314
	Cameron[7]	AusEng	230
	gymnast stefie	AusEng	277
	holly	BrEng	479
	Heidsta xxx	AusEng	398
			2106 posts
			Total: 5493 posts

Nevertheless, a more serious problem in Rank 2 needed to be solved. Both Rank 2 members *Jonathan* and *AnnaBessonovaNumber1Fan* are experienced users whose total number of posts far exceeds 2000. It posed great difficulties for the author of the present study to find another Rank 2 member of Australian or British heritage for an equal number of members in Ranks 1 and 2. Since a number of native speakers of British and Australian English had contributed between 900—1000 messages to the board, the retrieval of 750 posts from their profiles would go beyond the limits set in Table 3. The decision was made to extract only 242 posts from the profile of *RGribbonqueen* (with a total number of 942 messages), which still classifies the member as a Rank 2 user.

The posts are retrieved from each user's profile, where all messages are listed in chronological order. Each page displays 15 messages at once. Since those 15 messages are displayed in a contracted form and the context in which each post occurs is crucial for their subsequent coding, all messages are entered one by one and the whole page with 15 posts

[6] AusEng = Australian English; BrEng = British English.
[7] Nickname abbreviated in consideration of anonymity.

including the message in question were copied and pasted into separate text files.

Step 3. The identification of utterances containing the epistemic think *and evidential* seem

Once the posts were stored in 12 text files, all instances of epistemic *think* and evidential *seem* were tracked and highlighted. They were then delivered with their immediate context as well as such extra-linguistic information as rank (both of the sender and addressee) and topic of discussion to another text file in which each line represents one example. Altogether, 729 examples of *think* and *seem* were identified (see Table 7). From the text file, the examples were transferred into the software Filemaker Pro 11 in which they were coded for 31 factors (see Chapter 5.2 for the coding manual).

Table 7. Summary of utterances with *think* and *seem*

Rank	***think***	***seem***	**Total**
RANK 1: moderators	228	74	302
RANK 2: hosts	159	29	188
RANK 3: casual senders	189	50	239
	576	**153**	**729**

As Table 7 indicates, the total number of *think* constructions is nearly four times higher than utterances with *seem*. In fact, in previous research, *think* has been found to be the most prototypical epistemic marker in naturally occurring language (Biber et al. 1999; Kärkkäinen 2003). However, the unbalanced representation of constructions framed by *think* and *seem* in the data is overcome by the adoption of multivariate statistics (see Chapter 6). Also, the aim of this study is to draw a unified picture of both types of stance markers. Therefore, both verbs are preserved for subsequent analyses.

In summary, the two distinct processes of data collection described above are crucial for the adoption of both qualitative and quantitative analyses of stance constructions. The analyses conducted in the next few chapters are largely dependent on the decisions made in the course of data retrieval. Thus, Chapters 4—6 describe the two methods, dialogical discourse analysis and multifactorial usage-feature analysis, and present the results obtained through the adoption of these complementary approaches to stance-taking.

Chapter 4

A dialogical approach to epistemic and evidential stance

4.1 Dialogical discourse analysis

In *Rethinking language, mind, and world dialogically,* Linell (2009: 377—384) proposes a few models for the investigation of dialogically constructed interaction. The author also argues against the general assumption that it is the method that makes the analysis dialogical, mainly due to little sensitivity for dialogicality in data, but claims that there are a few methods that capture the intersubjectivity and collectiveness of meaning-making.

Linell's first method is the initiative-response (IR) analysis. It is an approach to dialogically conceived language that was developed to capture and quantify the dominance of basic turns in interaction. The main goal of the method is to analyze each turn's relations to previous and possible next turn units, while largely ignoring their content. This means that the model presupposes the existence of interaction with frequent turn units and therefore regards Internet bulletin board entries as too long. However, Linell also notes that one can develop what he calls a 'dialogical discourse analysis'. As opposed to IR analyses, this method goes into content and captures the dynamics of themes, recurrent features and other internal and external factors of dialogically constructed communication.

Although it is not the method that makes the analysis dialogical, Linell has proposed a number of features that are more appropriate for studying dialogicality than others (2009: 382—383). Among those, the author has listed two elements present in the two methods described above: the sequential and the thematic representation of communication. In fact,

these aspects, Linell notes, "may well make up the basis of what we may call a 'dialogical discourse analysis'" (2009: 384). Therefore, the method adopted for the present study, dialogical discourse analysis, unifies the underlying properties from two methodological models established within the framework of dialogicality and attempts to account for the sequential and thematic organization of stance. Also, its purpose is to show that stance is a socially as well as cognitively constructed phenomenon that emerges in dialogical conversations between two or more co-participants.

As indicated in Chapter 3.2, extracts from two threads from rsg.net are analyzed using dialogical discourse analysis. The approach follows two main underlying ideas to the investigation of stance. The first is sequential and the second is thematic. The sequential dependency of adjacency pairs is typically considered to be the backbone of Conversation Analysis (Sacks, Schegloff, and Jefferson 1974), but sequential organization also plays a central role in dialogicality. Dialogicality assumes that every word is directed towards an answer and therefore every turn taken by a speaker or writer is related to prior as well as possible future turns and contributions (Bakhtin 1981: 280). As a result, the first perspective of the method encompasses the sequentiality of epistemic and evidential stance in larger textual units, in which stance predicates are studied relative to their position and alignment with other contributions in the two rsg.net threads. In the present study, the approach primarily attempts at answering the following question: To what extent does the conversational participant govern or control the contributions made by members of rsg.net, and in turn, how is he/she governed and controlled by them?

The second perspective indicates that besides the sequential organization of stance, dialogical discourse analysis also takes into consideration the thematic and topical patterns of discourse in which epistemic and evidential verbs are generated. In this study, I portray epistemic and evidential verbs as crucial elements in framing commitment and attitudes towards evidence. Each marker is assumed to lead the thematic orientation of interaction and play an important role in the construction of opinions, ideas and judgments towards the topics being discussed.

However, when investigating the behavior and interdependence of stance constructions in discourse, we need to bear in mind the research questions presented in Chapter 1 and their interest in the strength of epistemic and evidential stance verbs from three hierarchically different ranks. Therefore, the sequential and thematic organization of stance-taking

in bulletin board conversations is above all studied relative to the interaction between moderators, hosts and casual senders, and the various techniques used to show authority and tentativeness.

4.2 A dialogical analysis of two threads from rsg.net

In the following section, I present the analyses of three extracts (see the Appendix) retrieved from rsg.net conversations. The extracts are studied in terms of the interdependence of epistemic and evidential stance and social rank. The analysis attempts to answer the second research question presented in Chapter 1:

— How do members from three different ranks attenuate or reinforce the strength of their propositions relative to their conversational co-participants?

For easier navigation between the objects of the study, the stance constructions are highlighted and analyzed according to the criteria established in Section 4.1.

4.2.1 Extract 1: Over-scoring

The first extract called *Over-scoring* (Extract 1 in the Appendix) displays examples of dialogically constructed epistemic and evidential stance markers by users from two different ranks: 2 and 3. It includes the initial message posted by the Rank 3 user *Brivido* and posts directly addressed to its content. The opening question — *Why does everybody hate Aleksandra Merkulova?* — is elaborated further by two users from a higher rank. As a result, six posts are represented in the extract. The protagonist of the discussion is a rising Russian gymnast Alexandra Merkulova, whose popularity in the community has been jeopardized by the supposedly corrupt and Russian-favoring judging system, or as rsg.net members like to call it, over-scoring. Lines 1—3 display the first post to the topic initiated by *Brivido*, whose inquiry about Merkulova's unfavorable position and portrayal in the gymnastics community serves as a trigger for subsequent messages. The user has framed his/her question in a way that encourages

other members to express their views and opinions on the topic, making the thread a promising object for investigating stance.

Epistemic and evidential stance first becomes apparent in a post by the Rank 2 user *Aleksandrafan* on lines 5, 8 and 13. *Aleksandrafan*'s message includes two instances of epistemic mental predicates (*I know*, *I think*) and one marker of evidentiality (*I remember*). In each case, the subject of the act of stance-taking is the poster, who positions himself/herself as 'who' in the discourse by using the first person singular marker *I*. As a result, the member takes advantage of three different stance constructions, *I know, I remember* and *I think*, to demonstrate his/her commitment to the propositions they modify. The first of these, *I know*, precludes the proposition *a lot of people don't like her style* on lines 5—6. Although it is the first instance of stance in the thread, this construction is particularly interesting from the dialogical point of view. Dialogical discourse analysis assumes that speakers and writers are not autonomous individuals operating in enclosed capsules, but rather social persons active in the intricate web of meaning-making. It is believed that their contributions are interdependent with previous as well as following contributions provided by others (Marková et al. 2007). Therefore, *Aleksandrafan*'s first epistemic stance construction can be interpreted in terms of the dialogical apprehension of possible subsequent messages, where the user shows his/her awareness of people's hostility towards the gymnast. Moreover, the user indirectly addresses these people and does so using rather bold generalizations (*immature people*, *a lot of*, etc.). Another important epistemic marker in *Aleksandrafan*'s post and arguably the most typical stance marker in the English language is the mental predicate *I think* (line 13). Here the user takes a strong evaluative stance, as opposed to the previous mental predicate, where *Aleksandrafan* rather reflected on the general opinion of the board than his/her viewpoints. *I think* can be considered as a continuation of an objective construal of public opinion to which the user provides an evaluative stance. As a result, the proposition the mental predicate *I think* modifies on line 13 (*is stupid*) exhibits the writer's subjective view on the general opinions being expressed on rsg.net about Aleksandra Merkulova. By doing so, *Aleksandrafan* explicitly attacks the users of the board and does so without any mitigating or softening devices to avoid subsequent imposition.

The initiator of the thread, *Brivido*, returns on lines 19—24, where he/she exclusively addresses *Aleksandrafan* and compliments the latter's reasoning skills (*Ah what a pleasure to read you*). Throughout the whole

post *Brivido* elaborates further on *Aleksandrafan*'s ideas and aligns his/her preferences with that of the other (*Like you, I don't like gymnasts like Miteva, Staniouta*). It is not until *Brivido* proposes a new idea on lines 23—24 that a mental predicate with an attempt to hedge its content is produced. The reason for this might lie in the novel representation of an idea, which serves as *Brivido*'s first explicit verbalization of an opinion, since until now the user has managed to avoid positioning himself/herself relative to the issue. What is more, the mental predicate in this case succeeds rather than precedes the proposition it modifies, a technique not yet utilized in the thread. In Holmes (1990) and Aijmer (1997), the authors have assigned a two-way classification to *I think*, deliberative and tentative, where the first expresses assertion and reassurance and the second uncertainty and mitigation. According to them, a predicate is deliberative when it has prosodic prominence, it is followed by the *that*-complementizer and occurs in a clause-initial position, while the tentative *I think* possesses opposite values, including being positioned in clause-medial or clause-final slots. In the present study and especially in the quantitative analysis of epistemic and evidential verbs (see Chapters 5—6), such a strict classification has been discarded at the expense of other contextual implications. However, in the present case these findings give some weight to what is being implied by the content of the proposition *I think* modifies: *Brivido*'s inclusion of the clause-final *I think* exhibits a great deal of tentativeness in terms of thematic composition and strategic positioning.

The importance of previous stance-taking in discourse becomes apparent when we look at the next turns. On lines 25—28, *Aleksandrafan* returns to answer the stance taken by *Brivido* about Merkulova's resemblance with a former Olympic Champion Alina Kabaeva, a gymnast whose energetic and youthful style was a turning point in the 1990s gymnastics. As the two evidential constructions on lines 25 and 26—27 are a reply to the previously investigated stance posted by *Brivido* on line 24, it functions as the first instance of intersubjective alignment between two co-participants. Intersubjectivity here follows the definition put forward by Du Bois (2007: 159), in which intersubjectivity is the result of a socio-cognitive relation that emerges between two subjectivities or when one speaker's subjectivity reacts to another's. Therefore, cases where a stance marker is in close affiliation with a previously occurring construction provided by another co-participant in the dialogical situation display intersubjective alignment of stance-taking. In both instances, *Aleksandrafan* uses the evidential *see* as a linguistic device to disagree with

Brivido's comment. This is indicated by the negative form of the verb *see* on lines 25 and 26—27 (*I don't see*) and the position adopted on line 28 (*I would rather...*). What is more, the verb constructions are produced without the presence of any mitigating devices, such as hedges or non-intensifying adverbs, and therefore display a rather strong and assertive content.

Table 8. Stance alignment between *Aleksandrafan* and **Kalinka**

	Username	**Evidential verb clause**	**Proposition**
1.	*Aleksandrafan*	I don't see	Kabaeva having charisma.
2.		I don't see	many similarities between Kabaeva and Sasha.
3.	**Kalinka**	I neither find	Kabaeva charismatic [...]
4.		I find	Merkulova very talented and charismatic [...]

Intersubjective dialogue is elaborated further in the post of the next member, **Kalinka** from Rank 2. With the evidential *find* on lines 34—37, **Kalinka** poses an explicit evaluative statement which is at the same time targeted at *Brivido*'s post about the similarities between the two gymnasts as well as *Aleksandrafan*'s opposing views on the matter. Besides an obvious alignment of stance between *Brivido* and *Aleksandrafan*, another and more fine-grained dimension can be detected in the messages by *Aleksandrafan* and **Kalinka**: the compatibility of evidential verbs by the two users (see Table 8, above). While *Aleksandrafan* contrasts his/her views with those of *Brivido*'s with *I don't see Kabaeva having charisma* and *I don't see many similarities between Kabaeva and Sasha* (lines 25—28), **Kalinka** achieves the same effect with *find* (*I neither find Kabaeva charismatic* on line 34 and *Whereas I find Merkulova very talented and charismatic* on lines 36—37). The repetitive use of the same number (two) and type (evidential) of stance markers might imply that **Kalinka**'s contribution in the dialogical interaction is a direct result of *Aleksandrafan*'s. This demonstrates that stance is the result of intersubjective relationship dynamically created by participants in the ongoing discussion. What is more, the first stance construction by both users (examples 1 and 3 in Table 8) modifies a proposition that addresses the same issue, a comparison between Kabaeva and Merkulova, while the second (examples 2 and 4) shows preference for Merkulova. Both utterances also leave an impression of firm and strong beliefs in one's

statements and can therefore be considered as instances of authoritative stance.

In the last message of extract (1), the original poster *Brivido* draws conclusions based on the opinions expressed by rsg.net members thus far. The message takes excessive advantage of epistemic mental predicates, showing that being challenged by two strong members from a higher rank (*Aleksandrafan* and **Kalinka**) impels one to display tentativeness and express likelihood rather than certainty and high possibility. The primary device for the marking of stance in this post is *I think* on lines 60, 65 and 71. The first *think* on line 60 displays a relatively high degree of uncertainty, but little chance for others to disagree with (*Merkulova is the little dog of Viner; I think that's right*). The second epistemic verb on line 65 precedes a proposition that would receive exhorting replies from the majority of the board if not for the subsequently inserted conjunction *but* and the proposition *she's not as clean as Charakashyna and Miteva* in the subsequent clause. This technique partly cancels the proposition in the previous clause and adds to the negative evaluation of Merkulova. With the last *I think* on line 71, the user first expresses sympathy towards Merkulova and her charismatic performances before cancelling the stance with *but youngness & innocence are not eternal* (lines 73—74). By doing so, *Brivido* minimizes the argumentative value of his/her attitude towards the importance of youthful energy in gymnastics. As a result, *Brivido*'s stance is marked as highly tentative and unlikely to trigger negative responses from other members. It can be assumed that the perspectives adopted by the user in his/her last post is prompted by the Rank 2 users, who despite their authoritative style display a certain degree of respect and attempt to maintain the addressee's positive face, which is the opposite of what can be found in the next extract.

4.2.2 Extract 2: Maybe you are blind

The following extract (Extract 2 in the Appendix) is extracted from the same thread as extract (1) and represents a subtopic developed in the course of the thread *Why does Merkulova receive so many critics?* This is explained by the fact that bulletin board threads are prone to exhibiting a great deal of fragmentation and display the emergence of multiple conversations and multiple involvement of users in several conversations (Marcoccia 2004: 120). Extract (2) represents a number of

counterarguments directed at the last post of extract (1). More specifically, it addresses one particular sentence: *I don't know, maybe I'm always optimist, but I can't imagine judges being corrupted*. The example is one of the conclusions drawn by the Rank 3 user *Brivido*, who does not seem to agree with previous comments on judges' preference of Merkulova over other gymnasts. This is a view that contradicts other attitudes prevalent in rsg.net. The extract starts with a comment made by a Rank 2 user, who directly addresses *Brivido*'s opinion on the judging system in rhythmic gymnastics, and as a result, introduces a new subtopic.

The first post on lines 76—77 is a sarcastic and rather condescending reply to *Brivido* by an experienced and expressive Rank 2 user *ybalka_*. It contains an imitation of laughter followed by the proposition *I would love to live in your bubble! Brivido*, as the member addressed, immediately provides a counterattack on lines 78—84. The post contains two markers of epistemic stance: *I don't think* on lines 78—79 and its positive form *I think* on line 82. In the first case, the poster re-emphasizes his/her earlier statement about corruption in rhythmic gymnastics (*No... really, I don't think judges are corrupted!*). The absence of an epistemic marker in an earlier post has been replaced by a construction that displays a high degree of certainty and strong evaluation by the writer. This is mainly exemplified by the presence of the intensifying and truth-attesting adverb *really* as well as an exclamation mark used to express emphasis. *Brivido*, whose previous post contained a number of devices for minimizing the argumentative value of his/her message, shows a great deal of brutality after being attacked by a Rank 2 user and uses *I think* to express reassurance rather than uncertainty. This shows that stance is situated in discourse and emerges in collectively constructed settings. In addition, the meaning of epistemic and evidential verbs can be highly flexible and context-dependent by obtaining meaning relative to users' conversational co-participants. The second epistemic marker on line 82 is another example of *I think* (*Howerer.. I think judges are always guided by their unconscious, like every human being*). In this case, the mental predicate is again accompanied by an adverb (*always*) that gives weight to *Brivido's* position and confirms the user's beliefs on the matter.

The last post in extract (1) represented on lines 92—112 is the first example of a message written by a Rank 1 user. Although moderators are omnipresent in the discussions unfolding in rsg.net by acting as 'guards' of the board, their contributions tend to be infrequent and sporadic due to their small number on the board. Therefore, the extract serves as an important

indication of the use of epistemic and evidential verbs by Rank 1 members. The extract, however, shows that the Rank 1 user *Storm* does not use any verbs exhibiting speaker commitment or source of evidence. *Storm*'s reply is resolute and even ruthless, starting with a long row of exclamation and question marks and a continuous use of words with capital letters (e.g. *YOU, RUN, EASY, UNFAIR*). In addition, the post exhibits propositions that show no signs of mitigation by epistemic and evidential markers. It is obvious that the message is produced with a deluge of emotions and strong opposing feelings towards *Brivido's* statements, but nevertheless, the inclusion of epistemic and evidential verbs to enhance conviction and reassurance has been discarded from the message.

4.2.3 Extract 3: It's never too late

The third and last extract (see Extract 3 in the Appendix) deals with a request posed by the Rank 3 user *OlympianRG,* whose dream to become an Olympic gymnast triggers an intense discussion in the thread *Can I still become an Olympian at age 19?* The thread contains six messages with lengthier replies than encountered in previous extracts. With longer messages, more epistemic and evidential verbs are prone to occur, which makes the extract a resourceful ground for the study of stance in dialogical discourse. The initial message posted by *OlympianRG* acts as a trigger for subsequent messages, in which the member asks for advice from his/her fellow members about becoming an Olympian after having started gymnastics trainings at 19. Since most professional gymnasts start training as early as 3—4 years old, the member's request is prone to generate a number of conflicting replies.

The opening post by *OlympianRG* contains four epistemic mental predicates of which three represent the lexeme *know.* As opposed to its negative equivalent *I don't know, I know* does not seem to appear in spoken discourse as often as the former (28 as opposed to 5 in Kärkkäinen (2003: 37)). This member modifies his/her propositions with the epistemic marker that typically conveys a high degree of commitment or even acts as a 'factive predicator' (Palmer [1986] 2001: 11). Moreover, when one looks at the propositions modified by the lexeme on lines 113, 119 and 136, a certain pattern tends to emerge (underlined):

(1) *Ok, I know it sounds old to start RG at age 19, but I am really determined and prepared to work hard.*
(2) *Please don't be confused by the 'Oympian' in the title, I know aiming for gold might not be possible, but just getting in the Olympics would already be heaven for me.*
(3) *I know it might take years and years to get to my goal, but I'm not a quitter and am prepared to work very hard.*

The mental predicates in all three examples modify propositions that first make indications about the futile probability of the dream, followed by the user's confirmation of the seriousness of his/her plans to become an Olympic gymnast. The clause modified by the mental predicate, therefore, seems to represent popular views present on the board, which are complemented with the member's own perspective on the matter. As a result, *OlympianRG* uses *I know* as a frame for propositions that represent generally accepted attitudes in the community, a device that significantly minimizes the risk of losing one's face later in the discussion. These examples show that the user makes way for possible new turns by pointing forward in the discourse and acknowledging the absurdity of his/her dream. From a dialogical point of view, *OlympianRG*'s use of the epistemic verb supports the idea of humans as social beings and co-authors of each other's contributions (Linell 2009: 73).

The first reply in the extract is a message from the Rank 3 user *Sasta33,* who provides a long answer on lines 140—204. The tone of *Sasta33's* message is generally positive towards *OlympianRG's* request, as exemplified by the following line: *I know there's going to be some people who will say that 19 is a bit old to start, but I am not of those people* (lines 140—142). As a result, the member takes a stance towards the opening post and aligns his/her usage of the epistemic mental predicate *I know* with that of *OlympianRG*'s. As scrutinized in the previous paragraphs, the latter also modified a similar proposition with the epistemic *I know*. In both cases, the clause being modified is representative of the general values of rsg.net, while what follows embodies the users' own judgments on the matter. Similar to *OlympianRG*, *Sasta33* takes an opposing stand to the general view and establishes an intersubjective alignment of stance-taking with the previous user. The following stance markers are mainly representations of the epistemic mental predicate *I think*, in which case the verbs are produced to display pure personal opinions and subjective impressions based on

one's internal state of mind (lines 142, 145 and 174). However, no significant indications of inter-user or intra-user alignment can be found.

The three following examples of epistemic and evidential alignment are portrayed in the posts by three rsg.net users: two members from Rank 2 and one member from Rank 1. The three utterances share a number of similarities in terms of the stance predicate *I think*, the presence of the second person *you/your* in the subject position of the complement clause, and a negative attitude towards the problem posed in the opening message. What is more, as the discussion evolves, the same epistemic verb is used to give increasing strength to the content of the proposition it modifies. This means that while the first post shows some awareness of *OlympianRG*'s face wants, the last message discards all devices for showing respect for the fellow member.

The rsg.net user **Kalinka**, who was also an active participant in previous extracts, provides his/her reply on lines 205—227. The epistemic and evidential markers are represented on lines 205, 207 and 214. However, only the second example framed by the epistemic mental predicate *I don't think* will be investigated here. In Kärkkäinen's classification of the functions of *I think* in pre-and post-positional slots, the example can be categorized as what the author refers to as the 'recipient-oriented design of utterances' (Kärkkäinen 2003: 146). These stance markers typically occur in more demanding trouble spots in interaction, where speakers and writers have to design and redesign their utterances to adjust to the characteristics of the ongoing discourse. The Rank 2 user **Kalinka** hedges his/her proposition with the mental predicate *I don't think* and provides the first example of a user who does not consider the dream to be achievable (lines 207—209). However, the member tries to remain on friendly terms with *OlympianRG*, which is exemplified by an encouraging assertion preceding the judgment on lines 205—206 (*It's never late for practicing the sport you love*) and a device for showing awareness of *OlympianRG*'s face wants on lines 206—207 (*I don't want to be cruel*).

On lines 228—232, the user *Invisible Hedgehog* from Rank 2 provides his/her comment on *OlympianRG*'s request. *Invisible Hedgehog* is a well-known user, who due to her strong affinity with the Ukrainian National Team possesses valuable information about the gymnastics world. As a result, the user can be described as a respected member in the community, whose favorable position on the board as well as in the gymnastics world outside it is recognized by many. With some background knowledge of the

tendency for gymnasts from highly competitive countries to move to locations where qualifying for title championships is somewhat easier, one is confident to say that *Invisible Hedgehog*'s suggestion framed by the epistemic *I think* (line 228) displays a case of 'irony'. Much research has been conducted on the nature of ironic utterances (for example, Sperber and Wilson 1981; Gibbs and O'Brien 1991; Giora 1995) and the basic functions these devices have in language. For instance, Giora (1995) has proposed irony to be a form of indirect negation through which it is expressed without an overt negative marker. It can therefore be assumed that the reply posed by *Invisible Hedgehog* is produced to implicitly suggest the absurdness of *OlympianRG*'s request. As a result, *Invisible Hedgehog* rather tactfully opposes herself to the issue by tackling on the sensitive problem present in rhythmic gymnastics.

No instances of politeness strategies or implicitness are present in a post by the Rank 3 user *uscoach* on lines 233—247. It is first implied in the opening sentence of the message on line 233 (*Simple answer "NO"*) and followed by a thorough overview of the reasons behind *uscoach*'s stand. Epistemic commitment is first presented on lines 245—246 with the mental predicate *I think* and accompanied by an adverb (*honestly*) and a number of demoralizing adjectives, such as *ridiculous* and *disrespectful*. Therefore, a number of strong discursive features indicate the verb's assertive and authoritative meaning. What is more, the tendency for the epistemic stance construction to grow in strength while being directed at the same issue, using the same mental predicate and addressing the same person, suggests that the three examples are produced by building on previous turns. Although all three users implicitly acknowledge the absurdity of the request, not everybody states it explicitly. The first example by the Rank 2 user has been designed with an aim to stay on friendly terms with *OlympianRG.* The second makes use of irony, which at the same time negates the possibility of the request. However, these strategies do not seem to satisfy the Rank 3 user *uscoach*, whose proposition framed by *I think* is used with absolute authority and enhanced by various contextual cues.

For the extract to be representative of all three ranks operationalized in Chapter 3.1.3, an entry posted by the Rank 1 user *Tahnee* has also been included in the extract. This is the only post by a Rank 1 user in the thread. However, on lines 248—273, where *Tahnee* gives a rather profound explanation as to why starting at age 19 and going to the Olympic Games is not a reasonable idea, no epistemic and evidential markers can be found. As

a result, *Tahnee's* opinion and judgment is not framed by devices that empower or mitigate one's propositions, but nevertheless displays a resolute and tenacious answer to *OlympianRG*'s plea for help. This is also in line with another post by a Rank 1 user in extract (2).

4.3 Discussion

The dialogical discourse analysis adopted for the investigation of three extracts from rsg.net needs to account for a number of communicative and conceptual features of stance-taking. First, it needs to exhibit the sequential organization of epistemic and evidential verbs. Secondly, the method needs to account for the thematic progression of topics and sequences framed by these stance markers. Thirdly, in accordance with the research questions presented in Chapter 1, it needs to answer the question of how rsg.net users align stance markers relative to their conversational partners in terms of social rank. The analysis of the three short extracts shows that rsg.net users often align their epistemic and evidential stance constructions with that of others, which shows that their construal of the world is strongly interdependent with their fellow members (discussed below as Result 1). In addition, despite the presence of a number of differences in the construction of stance by members from three hierarchically different social ranks, further analysis is crucial for the detection of authoritative and tentative stance production in bulletin board conversations (Result 2).

Result 1. Rsg.net members align their stance constructions with fellow users

The interdependence of stance is exemplified by a number of properties. First, it becomes clear that rsg.net members both react to each others' subjectivity as well as point forward in the discourse and make grounds for subsequent evaluations. The Rank 1 user *Brivido* from extract (1) is an example of such alignment. After *Brivido*'s tentative positioning of himself/herself relative to Merkulova's similarities with the former Olympic Champion Alina Kabaeva, this stance is directly addressed by two Rank 2 users: *Aleksandrafan* and **Kalinka**. Their successive responses with such evidential markers as *find* and *see* pose a clear example of alignment in terms of type (epistemic vs. evidential) and strength. The example shows an interesting alignment of stance and demonstrates the fine-grained nature of construal between different writers. *Aleksandrafan*

and **Kalinka**'s alignment of two evidential verbs that agree in terms of type, number, strength, and most importantly social rank, shows that people are constantly affected by the speech and writing of others. In the present case, this is exemplified by epistemic and evidential stance through which people evaluate, judge and make sense of the world around them. By tackling the same issue first presented in *Brivido's* message, the following contributors add their evidential constructions like perfectly fitting building blocks.

Another example of the interdependence of stance in the three extracts emerges from extract (3), where three users are involved in expressing attitudes towards propositions with similar content. The epistemic marker, *I think*, serves as the shared element of stance and in all cases modifies propositions that oppose to the idea of becoming an Olympic gymnast after having started gymnastics trainings at the age of 19. Similar to the metaphor of building blocks introduced above, the same can be applied to the present case. However, in extract (3), the building blocks are of varying sizes. This means that while the first member **Kalinka** starts with an utterance containing a number of mitigating devices and the second user takes advantage of irony to hide his/her negating view, the last member provides a rather brutal and discouraging reply to *OlympianRG*'s request. This might indicate that by being preceded by a number of indirect suggestions, the Rank 3 user *uscoach* produces a stance that exhibits a direct and assertive truth about the hardships of becoming an Olympic gymnast. *Uscoach*'s reply can therefore be considered a result of the unfolding conversation, in which previous answers do not seem to satisfy the member's views and attitudes.

The second implication of dialogically constructed language in rsg.net is what can be called as 'pointing forward in the discussion'. One of the theoretical issues present in the study of dialogicality is the embodiment of speakers and writers as social beings who are in constant dependence on each other's language, thought and experience. This means that dialogicality is not only concerned with reactions to previously occurring turns but also prepare the ground for subsequent speech acts. In most cases, this was achieved with the epistemic mental predicate *I know* that typically demonstrate the factitivity of the propositions they modify as well as high certainty towards their truth and reliability. However, the propositions framed by *I know* in these cases contained acknowledgements of the general views popular in rsg.net as opposed to subjective impressions on the part of the writer. By doing so, these members first show awareness of

public attitudes before introducing their deviating views on the matter. What is notable is that the majority of examples with *I know* in thread-initial positions have been posted by Rank 3 users (discussed in the next part).

Result 2. The representation of stance constructions across three different ranks shows negligible differences

So far, I have primarily demonstrated the analytical strength of dialogical discourse analysis through which stance can be studied. However, as the attenuation and reinforcement of stance relative to social rank acts as the main objective of the qualitative study, these findings need to be combined and studied in terms of social rank. The analysis of the three extracts shows that a few differences can be found in the techniques used by members from three different ranks. However, these results should be complemented with additional methodological tools for better understanding of the authoritative and tentative qualities of stance in bulletin board conversations.

The usage of epistemic and evidential stance by Rank 3 users shows a relatively high degree of variation. In addition to instability in inter-user communication, the techniques used also have minor differences within messages from the same member. However, perhaps the clearest feature becomes apparent in the usage of *I know* to modify propositions without an overt marking of subjectivity. With the adoption of the predicate, Rank 3 users typically tend to acknowledge the possible opposing views prone to emerge in the conversation before expressing their own perspective on the matter. To some extent, they also minimize the chances of being imposed on by other members of the board by acknowledging their views and attitudes. This shows that these members are aware of the argumentative and contentious nature of rsg.net threads and bulletin board conversations in general.

The example from the Rank 3 user *Brivido*, however, shows that stance is not only dependent on social rank but on the nature of input by other members of the board. In both extracts (1) and (2), the Rank 3 user *Brivido* addresses users from a higher rank, but the strength and nature of these posts differ significantly from one another. When closing extract (1) with the conclusions drawn from the messages provided by others, *Brivido* skillfully avoids taking a strong stand towards Aleksandra Merkulova's position in the gymnastics world. All propositions rather represent the views posed by others than *Brivido*'s own perspective on the topic. In

addition, all replies to *Brivido*'s initial inquiry show no or little degree of imposition and argumentativity. However, after being opposed by a Rank 2 user in extract (2) with a sarcastic comment on *Brivido's* viewpoint, the member uses the epistemic *think* and accompanies it with features that show strong commitment to his/her beliefs. This indicates that meaning and sense-making is co-constructed between members in conversational settings rather than display a fixed values depending on one's social rank.

When it comes to Rank 2 users, no distinct devices for the production of stance can be found in the data and differences between members from Ranks 2 and 3 are paper-thin. For instance, in the negotiation of stance in extract (3), two Rank 2 users tend to display a great degree of politeness and face saving strategies by avoiding being explicitly blunt, while the most brutal attack is provided by a Rank 3 user. Another contradiction can be found in extract (1), where the provider of the first reply, *Aleksandrafan*, uses the mental predicate *I know* to acknowledge the presence of contradicting views on the judging system in rhythmic gymnastics. The same technique was detected in the speech of Rank 3 users. However, as the message unfolds, the same Rank 2 member also verbally attacks the members whom he/she addresses and construes a negative opinion of them. At the same time, the Rank 3 user does not explicitly judge the views adopted by the people he/she has addressed.

Although there are few Rank 1 users in the three extracts, two posts were provided by two members from this privileged rank. Both members have held the position for a number of years and play an important role in the board's everyday life. What is interesting is the nonexistence of epistemic and evidential stance markers in both users' posts on such controversial topics. The absence of these devices, however, does not suggest that that the language used by Rank 1 users would in any way be impartial. In contrary, stance in these posts is expressed through other techniques available in Internet domains, such as the abundance of exclamation and question marks, the emphasis of key words with capital letters or simply the inclusion of emotional and affective content words. Therefore, subjectivity from these users is conveyed through other means, which deserve more attention in future research.

Since dialogicality is a qualitative phenomenon and dialogical discourse analysis a methodological framework developed within its theoretical boundaries, researchers working in the field strongly believe that interactive qualities in speech and writing cannot be reduced to quantitative measures (Marková et al. 2007: 28). The present case study has

shed light on the behavior of epistemic and evidential verbs in context. However, for a more representative overview of stance and the qualities it possesses to perform socio-functional purposes in human interaction, a more large-scale and quantitative study needs to be conducted. For this reason, Chapters 5 and 6 complement the results obtained from the present chapter by adopting a quantitative corpus-driven approach to stance-taking.

Chapter 5

A quantitative approach to epistemic and evidential stance

5.1 Multifactorial usage-feature analysis

The methodology of dialogical discourse analysis implemented in the previous chapter showed that stance constructions are an important tool for accomplishing intersubjectivity between members of rsg.net. It also demonstrated that qualitative analysis displays limitations in terms of the presence of epistemic and evidential stance constructions from all three ranks. This is primarily due to the relatively small sample retrieved from the board. In the following chapters, the previous study is complemented with an entirely different perspective. I adopt a quantitative corpus-driven approach to stance-taking. More specifically, the research adopts the multifactorial usage-feature analysis to study the structuring of epistemic and evidential verbs. Adopting this method, I attempt to answer the third research question presented in Chapter 1:

— What formal, semantic and extra-linguistic factors, framed by *think* and *seem*, indicate social hierarchy between the three ranks?

Multifactorial usage-feature analysis combines two important methodological viewpoints inherent in its name: usage-feature and multifactorial. To start with, usage-based linguistics encourages linguists to use data from naturally occurring language. It also assumes that language usage structures grammar, which in turn motivates meaning (Lakoff 1987; Langacker 1987). This suggests that by looking at linguistic patterns and their frequency of occurrence in real language, it is possible to explore the dynamic shape of language structure, and in line with the motivations

behind Cognitive Linguistics, its conceptual system (Bybee 1985; Hopper 1987; Tomasello 2003).

Multifactorial analyses reflect the relationship and interaction between prosody, syntax, semantics and other sociolinguistic factors. Gries (2003) notes in his account on particle placement in English that monofactorial or even pair-wise oppositions do not explain the complex interaction between linguistic and social factors present in language production. Multifactorial modeling, therefore, acknowledges the omnipresence of such usage features as semantics, pragmatics, lexis, and syntax, and accounts for their simultaneous involvement in interaction (Glynn 2010a: 9).

One of the advantages of implementing quantitative methods in linguistic research is the employment of statistical tools. The current study takes advantage of both exploratory and confirmatory statistical analyses. First, the exploratory technique, Multiple Correspondence Analysis (Glynn in press), helps to identify patterns in data and shows the strength of association between factors. With confirmatory techniques, researchers are able to confirm the statistical significance and predictive power of the patterns found with exploratory tools. Therefore, it is only with confirmatory tools that our results can be fully verified. The confirmatory technique used in the present study is Binary Logistic Regression Analysis.

Before conducting a multifactorial analysis of *think* and *seem*, Section 5.2 describes the criteria established for the coding of 729 examples retrieved from rsg.net. The coding schema comprises a number of objectifiable or formal factors (5.2.1); however, the most informative factors in determining social rank are assumed to make references to the semantic or subjective content of the constructions instead (5.2.2). Previous research has argued for the employment of subjective phenomena in quantitative research (e.g. Glynn 2010a), and the present study also demonstrates that with a fixed coding schema and the adoption of both exploratory and confirmatory statistical tools, the investigation of epistemic and evidential stance is a feasible task.

5.2 Classification and coding of *think* and *seem*

"Coding is the process of segmenting natural and experimental events […] into discrete labeled chunks for purposes of counting, analysis, or statistical manipulation" (Scheibman 2002: 23). Scheibman continues by acknowledging the complex and indiscrete task of classifying grammatical and semantic units into analytically devisable elements. However, the operationalization of qualitative phenomena into quantifiable units is vital for conducting the analyses adopted in the present study. Coding categories are typically constructed by the researcher based on traditional linguistic theories and orientations, but can also emerge from the data at hand. In fact, it is advisable for analysts to build coding schemas around their research questions for more productive use of time (Lampert and Ervin-Tripp 1993). In line with the present research goals and research questions presented in Chapter 1, relevant categories that are expected to account for the multidimensionality of stance in social settings have been formulated.

The coding, or annotation, of 729 examples of *think* and *seem* was conducted in a database called Filemaker Pro 11. It is a program developed to facilitate the manipulation of a large number of examples and their systematic annotation for relevant values. Consequently, all instances of *think* and *seem* with their variations *thought, seems* and *seemed* were coded for 31 formal, semantic and extra-linguistic factors:

- *Formal factors*: construction of the utterance, position of the verb clause, time reference, mood (main and complement clause), negation, person (main and complement clause), modifier
- *Semantic factors*: Chafe's source of evidence, Krawczak and Glynn's epistemic type, Fortescue's classification, Langacker's subjectivity, Kärkkäinen's classification of stance functions, pragmatic intention, aspect, degree of epistemic commitment, Scheibman's adverb type, evaluation, verification, emphasis, argumentativity
- *Extra-linguistic factors*: social rank (sender and addressee), dialect, topic of discussion

All factors are assumed to contribute to the identification of authoritative and tentative realizations of the two stance constructions.

The following sections are an overview of all 31 factors with a detailed description of the coding criteria. The characteristic features

assigned to each factor are called *values*. The minimal analyzable unit in the present investigation is the *utterance*, including both the main clause in which the verb occurs as well as the complement clause modified by the predicate. However, often, larger contexts need to be taken into account and cues for coding a number of semantic factors need to be searched from surrounding utterances and messages.

5.2.1 Formal factors

Construction of the utterance

The construction of the utterance is studied relative to the main clause with *think* or *seem,* as well as the complement clause modified by the verb. In the present coding schema, complement clauses modifying *think* are divided into clauses containing the copula *be* for establishing relations between subject and object and clauses with a full verb in their predicative slots. Characteristically to such stance constructions as *think, believe* and *guess,* main and complement clauses can be connected by the *that*-complementizer, a feature that is often omitted in informal conversations and face-to-face interaction (Thompson and Mulac 1991). For instance, example (1) is coded as a clause comprising the full verb *buy* and bridged by the *that*-complementizer, while in (2), the verb is a copula and the *that*-complementizer is omitted between the two clauses.

(1) *I'm live in NZL but I think* ***that*** *nearly all of the girls here,* ***buy*** *there equipment from Amco...they are really good!!!*
(2) *I think ∅ this version of her ball* ***is*** *better than the 2008 one!*

In addition to full clauses, *think* can also take the *so*-adverbial as its complementing unit. For instance, in *I think so, so* is used anaphorically and refers to a previously occurring clause.

Seem-constructions, on the other hand, require a relatively different treatment from *think*. The only construction shared with *think* is the finite *it seems that*, which is characterized by both features found in *think*-constructions: the presence/absence of the *that*-complementizer and the type of verb found in the complement clause. Other constructions coded in clauses with *seem* are retrieved from Aijmer's (2009) study on evidentiality, where the author has listed a number of possible *seem*-constructions found in English: copular phrases (e.g. *This seems low*),

comparatives (e.g. *This seems like a logical explaination*), infinitive constructions (e.g. *Nobody seems to care*), and finally, subjective *it seems to X* constructions with or without a complementizer, which can be followed by complement clauses, copular phrases or comparatives.

Position of the verb clause
Epistemic and evidential verbs can either appear in initial, medial or final positions, relative to the propositions they modify. Coding for clause-initial and clause-final instances of *think* and *seem* is rather unproblematic, however, criteria for clause-medial coding has to be established. For example, it has been demonstrated that these instances usually preclude the adoption of the *that*-complementizer (Kaltenböck 2013: 10). Consider the following example:

(3) *Viner is now their named trainer and she is the one calling the shots, then Sergaeva (who **I think** is a very talented coach) is always standing in her shadow and having to agree with Viner.*

The example displays the mental predicate in the medial position, since the inclusion of *that* between the stance construction and the proposition it modifies would result in an ungrammatical sentence structure (**who I think that*).

Time reference
The marking of time is determined in the clause the verb occurs in and is coded for three values: *past* (e.g. *I really didn't think*), *present* (e.g. *Everything seems*) or *future* (e.g. *Things will seem*).

Mood
The grammatical mood of both the main and the complement clause is assigned three values. Stance verbs and complement clauses with the presence of a modal marker (e.g. *I would think*) are coded as *modals*. Moreover, *conditional mood* is assigned to examples where the action in the main clause can only happen if a certain condition is fulfilled, as exemplified in (4).

(4) *I think the results would have been different had certain judges been judging...*

Finally, the *indicative mood* is the unmarked value and assigned to all other cases.

Negation
Similar to grammatical mood, negation is also determined in both clauses. In the clauses, negation can be *absent*, *neutral* or *strong*. In distinguishing the latter values from each other, fixed criteria have to be established. As a result, the absence of any intensifying markers, such as adverbs, capital letters and others, determines the clause as *neutral*, as in (5), while the presence of these features classifies the negation as *strong,* as demonstrated in (6).

(5) ***I don't think*** *that half the time they deserve it, even if maybe they still deserve to win.*
(6) ***I really don't think*** *Santoni has the same level as them (Stefanescu and Savrayuk for example).*

Clause-final negative markers with an adverbial *so* are also coded as *strong*. The reason for this lies in their seemingly strong impact on the content of the proposition, as exemplified in example (7).

(7) *Rhythmic steps during those 2 rolls on the body??* ***I don't think so****, not rhythmical at all.*

Person
The categorization of grammatical person varies largely between the two verbs. The factor is determined both in the clause in which the verb occurs as well as in the complement clause it modifies. In case of the former, all instances of the epistemic *think* in the present study always take the first person singular in its subject position and therefore show no variation in the coding schema. *Seem,* on the other hand, has a greater variety of possible values. It can adopt *first*, *second*, and *third person* subjects, and such *non-referential* subjects as *there* (e.g. *There seems to*), as well as the dummy pronoun *it* (e.g. *It seems that*). Additionally, third person singular and plural constructions are divided into two subcategories depending on their *human* or *non-human* referents, as demonstrated in the following examples:

(8) *I love how she expresses the music and **she** seemed confident with the handling which was a rarity in that competition.*
(9) *It goes without saying that **competitions** will seem a little empty without one of Russia's only contenders.*

The same criteria apply for complement clauses. Again, coding for complement clauses followed by *think* does not pose any difficulties; for example, *she* in example (10) is unanimously coded as third person singular with a human referent.

(10) *I don't think **she** will be going to the Olympics.*

The same classification is applied to *seem*-constructions with complement clauses similar to (10), such as *it seems that* as well as the subjective *it seems to X*. However, problems arise with copular *seem*-constructions that combine with a large variety of nominal and adjectival phrases without overt subjects. Therefore, examples like (11), where *seem* is accompanied by an adjective, lacks a subject in its complement phrase and refers to comments made on appearance or direct perception, are coded *N/A*. An identical treatment is applied to infinitive constructions with adjectival and predicative phrases, as seen in (12).

(11) *It seems **a little unfair** if Africa receives 2 wildcards (EGY plus CPV for the tripartite) and Oceania gets nothing...*
(12) *Kondakova wasn't having the best day and by the end, seemed to be **very upset.***

Modifier
This factor is coded for the presence or absence of an adverb that precedes the verb and is coded further for its type below (see *Scheibman's adverb type*).

5.2.2 Semantic factors

As opposed to the previous factors that were coded in a rather straightforward way, the following factors are considered to be more subjective. Although fixed criteria have been established, the manual coding for these factors is prone to a high degree of subjectivity and

variation, and is purely based on the intuitions of a single coder. For this reason, the adoption of confirmatory statistics is vital. In addition, for future research, multiple coders are included in the coding of the factors to permit the adoption of Cohen's Kappa score and determine the degree of inter-user agreement.

Chafe's source of evidence

As briefly described in Chapter 2, in Chafe's (1986) account on the marking of evidentiality in English, the author defines the term evidentiality in its broadest sense. According to him, the phenomenon does not only refer to the source of knowledge but also to the attitudes speakers and writers have towards it. Therefore, such prototypical epistemic mental predicates as *think* can also be found in the categorization of various modes of evidence. The modes of knowing listed by Chafe include *belief*, *induction*, *hearsay* and *deduction*. While *belief* often downgrades the importance of evidence, as shown in (13), the source of knowing is strongly present in the others. For instance, *induction* implies that some sort of inference in the speaker's mind has taken place by relying on evidence acquired though various channels of information transmission, as in (14). Additionally, example (15) or *hearsay* refers to knowledge acquired through language and *deduction* predicts what will count as evidence (no instances found in the corpus). Although not listed as a mode of knowing, Chafe introduces evidence acquired through *sensory* or perceptual channels, which is exemplified in (16) and treated as one of the values in the present coding schema.

(13) *Are taxis hard to come by in this area?* ***I think*** *I would rather get a taxi at that time of night than a bus.*. [Belief]
(14) ***It seems to me*** *as if she's accepted that this Olympics is not going to be fair, and just wants to do her best.* [Induction]
(15) *Ledoux with bronze* ***I think****... if I understood what Vera Atkinson said correctly.* [Hearsay]
(16) ***She seemed*** *rushed at the end. On her standing leap she just had the Rope piled into her hand.* [Sensory]

The present study also adopts a similarly broad approach to evidentiality as Chafe's, and argues that just like *seem* can simultaneously express evidentiality and epistemic necessity (Aijmer 2009), with the help of

enough contextual cues, *I think* as a prototypical epistemic marker, can also carry indications of source.

Krawczak's and Glynn's epistemic type

In their study on epistemic mental predicates in online weblogs, Krawczak and Glynn (2011) have coded four mental predicates for what they call the epistemic type. The factor includes the following levels: *belief, opinion, conviction, prediction, trust, confidence, activity, intention, request, estimation* and *idea*. For the coding of *think* and *seem*, the number of possible values is more restricted as the verbs show a more limited range of epistemic functions. The epistemic types found in rsg.net are: *opinion, conviction, intention, prediction* and *question. Opinion,* as demonstrated in example (17), is the unmarked value and contains instances where writers express their 'pure opinion' on a specific subject matter.

(17) *Ok, since it seems there are different opinions, I decided to open a topic for it.*

In addition, *conviction* is typically exemplified by such emphatic markers as charged content words and (the abundance of) exclamation marks, as in (18).

(18) *There's no wow factor, I think 2012 is Kondakova's year!*

Both *intention* (19) and *prediction* (20) make statements about the future. *Intention* reflects the agent's aim as to how certain situations are to unfold and *prediction* indicates the future outcome of events based on evidence.

(19) *I think i'll vote for AZE.... again I voted for AZE last year also!*
(20) *i think that next year i will be in Paris when its on!*

Only a few instances of the final function, *question,* were found in the corpus, but the occurrence of these demonstrates the wide range of functional possibilities of the two stance constructions:

(21) *Am I the only one to think she seemed pretty "stony"?*

Fortescue's classification

This factor in the coding schema follows Fortescue's (2001) division of the possible senses of *I think*. The mental activity 'thinking' is one of the core concepts studied by Fortescue and the lexical field the polysemous predicate *think* covers is syntactically and semantically broad and rich. Fortescue assigns the mental predicate three senses. *Think* as an 'unspecified mental activity', as in *think about,* is excluded from the coding schema due to our interest with epistemic meanings only. Therefore, only the two remaining values, *belief* and *evaluation,* are applied to the examples found in rsg.net. What distinguishes these two senses is the former's reflection of the propositional content of the message, as in (22), while the latter is concerned with its judgmental and evaluative value, as exemplified in (23).

(22) *When alina dropped the hoop at the end I cried so loud I think I woke up the guy next door.*
(23) *I think it's better to stay in the centre and take the bus there daily*

As an attempt to integrate traditionally evidential perception verbs into the investigation of epistemic modality, all instances of *seem* are also coded for the two senses originally intended for the analysis of *think* only.

Langacker's subjectivity

As opposed to the Functional Linguistics identification of linguistic patterns in text, Langacker's distinction between subjectivity and objectivity relies on purely cognitive properties. Langacker (1991: 315—342) describes subjectivity through the concept of visual construal and the relationship between the conceptualizer and the concept being perceived. The terms Langacker uses for the notions of subjectivity and objectivity are described by theater metaphors, *offstage* and *onstage* respectively. When a conceptualizer remains non-salient and implicit with respect to the object of perception, he/she is maximally subjective or *offstage,* while a conceptualizer who is part of the conceptualization, i.e. he/she is *onstage*, is maximally objective. Although most examples coded in the data rely on the objective relationship between the rsg.net member and the object being described, and *onstage* is classified as the unmarked value, there are a few instances where the stance construction acts as a perceptual apparatus through which the writer construes the situation at hand. For instance, the

following example has been coded as *subjective* due to implicit status of the conceptualizer with respect to what is being construed:

(24) *Dear katerin, Thank you for reuploading the videos but they dont seem to work!*

Kärkkäinen's classification of stance functions

In her work on epistemic stance in English conversations with an emphasis on *I think,* Kärkkäinen (2003) notes that in addition to marking deliberation and tentativeness as established in Holmes (1990) and Aijmer (1997), *I think* does much more functional and organizational work in interactive discourse. Furthermore, Kärkkäinen emphasizes the importance of *I think* in various positional slots.

First, pre-positioned *I think* can act as a *boundary marker* and introduce a new frame in conversations by "pointing forward in the discourse" (Kärkkäinen 2003: 128). Such instances are typically objectively construed and minimized for their evaluative value. Consider example (25).

(25) *I think there is a problem with the SUI group names =P*

Secondly, *I think* can have another starting-point function, which attends to some trouble in previous turns, and typically occurs in answers to questions, weak agreements and other second parts of adjacency pairs. The function is coded *speaker perspective* and exemplified in (26).

(26) *I think it's weird too, having EF before the AA*

Last but not least, *I think* can occur in more serious trouble spots in interaction, where writers have to design and redesign their utterances with attention to their conversational partners. This is coded *recipient oriented*:

(27) *Don't get me wrong, I think Chrystalleni is a very good gymnast - she clearly has the flexibility over Naazmi.*

Mental predicates occurring in post-positional or clause-final slots have been assigned one functional role, namely the signaling of *turn completion* on behalf of the writer of the post or the *uptake of response* from the interlocutor. Since sending a message into the Internet domain is the only

indication of a possible turn completion in bulletin boards (due to the lack of intonation cues), only stance constructions occurring in final positions of rsg.net messages are coded as post-positional.

Similar to Fortescue's classification of senses assigned to *I think,* Kärkkäinen's findings also apply to the epistemic verb only. However, since our treatment attempts at incorporating epistemic and evidential phenomena into a single coding schema, *seem* is coded according to the same criteria as *I think.*

Pragmatic intention

What speakers and writers do with language is determined by pragmatic acts of speech. In fact, without the illocutionary force of an utterance communication cannot even exist (Cruse 2004: 365). The posts written by rsg.net members convey various messages framed by epistemic and evidential markers that serve different pragmatic functions. The pragmatic intentions in these messages are identified and coded for the following values: *humor, praise, downtoning, prominence, complaint* and *insult.* To start with, *humor* entails both positive and negative comments about humans, events and other states of affairs. An example of positive humor is demonstrated in example (28), where the user touches upon the difficulties in handling the gymnastics apparatus ball.

(28) *[…] but also ball because always seems to find a way of running away from me! I think my ball grows legs.*

Praise is a common tool to show one's appreciation of gymnasts and groups and it typically manifests itself through the use of emotionally charged content words:

(29) *i think naazmi is a beautiful gymnast...she such a good character too...i judge with her and shes the funniest girl...*

Complaint and *insult* are opposite values of *praise.* While the object of complaint is typically a person outside the rsg.net community, such as professional gymnasts, coaches and judges, as in (30), *insult* is usually directed at other members in the community, as demonstrated in (31).

(30) *The Russians were great, although all seemed to lack speed in their routines.*

(31) *Youandi you can stop NOW with this wingy tune cause I don;t think there's such a HUGE chaos of spamming topics left around.*

The two remaining values, *downtoning* and *prominence*, are primarily concerned with politeness strategies. *Downtoning* is used to diminish the chances of threatening the face wants of the person being addressed. It is often recognized by the adoption of *we* as opposed to *you* in addressing critical issues or correcting interlocutor's propositions, while *prominence* tends to lack such characteristic features:

(32) *Maksymenko's routines may seem 'Boring,' (Which I do not agree with - She interprets the music) […]*
(33) *I think, unless you are put under a very intense training regime (I'm talking maybe girls in Russia), your bones will end up completely normal and you will be able to walk fine.*

Aspect

In Vendler's (1957: 143) much-cited article on possible time schemata indicated by English verbs, the author demonstrates how verbs presuppose and involve the notion of aspect. More specifically, Vendler proposes to divide verb aspects into four groups: *states, activities, achievements* and *accomplishments.* The first group, *states*, contains meanings lacking continuity and an inherent progression in time, such as *love, hate, know,* and others. Another verb type that lacks continuity is *achievement,* which differs from *states* with regard to its instantaneous nature. Vendler brings such examples as reaching a hilltop, winning a race, etc. Groups where verb phrases may possess a continuous tense include *activity* and *accomplishment.* The latter refers to events with an inherent ending point (running a mile, drawing a circle), while the former entails a continuous progression in time, where no assumptions about the duration of the activity are implied (running in general).

The predicates are coded in the complement clause following the two stance constructions. Examples (34), (35), (36) and (37) are instances of verbs indicating *state, activity, achievement* and *accomplishment* respectively.

(34) *Actually I think that RUS gymnasts* ***lack*** *sharpness in their movement.* [State]

(35) *I know when I went to LA Lights last year, it seemed that every coach **spoke** Russian - I could well have been in a foreign country!* [Activity]
(36) *I think the judges **have scored** those routines (Which were well performed, don't get me wrong) Too high.* [Achievement]
(37) *6. I don't think she **was using** the hoop as support, therefore I give her the 0.8* [Accomplishment]

It should be borne in mind that Vendler's classification of sentence types in the present study is only applied to complement clauses with full verbs. A different approach is applied to complement clauses where the subject and the non-verbal element are joined by the copular *be.* The reason for this lies in the fact that relational clauses are found to display examples where events, ideas and entities are assigned characteristics based on evaluations and speaker's point of view (Scheibman 2002: 158). The two aspects taken into consideration are the specificity of the subject, i.e. whether the subject is specifically singled out, and the inherentess of the quality applied to the subject (see Table 9 for examples). Although these examples are also considered to be *states* in Vendler's classification, only the relational affiliation of the subject and the non-verbal element is taken into consideration.

Table 9. Coding for the specificity and inherentness of copular clauses

	Inherent	**Non-Inherent**
Specific	I think **she** is **russian**, but she coached in south korea??	Yeah I think **purple** is **right** - 10kgs is a LOT and you need to be really careful, especially if you're not overweight to start with?
Non-Specific	I think that **every style** is **personal** and can't be imitated..	I don't think **an entirely "neutral jury"** is really **possible**, but it was a nice idea and I think the results would have been different had certain judges been judging..

Complement clauses that have the dummy subjects *there* or *it* in their nominal slots are coded as *abstract.*

Scheibman's adverb type

The criteria for coding adverb type are based on Scheibman (2002). Scheibman's categorization includes eight values, out of which two are excluded due to their incompatibility to modify epistemic and evidential verbs. An additional value, *other*, comprises adverb constructions not compatible with Scheibman's classification. The final arrangement with examples from rsg.net is presented in Table 10.

Table 10. Adverbial type codes (adapted from Scheibman 2002: 57)

Adverbial type	**Description**	**Examples**
Intensifier-amplifier	scale upwards from an assumed norm	even
Intensifier-downtowner	scale downwards from an assumed norm	just
Manner	means, quality, comparison	strangely, somehow
Modality	modification of the force or truth value of an utterance using emphasis, focus, or approximation	really, in fact, firmly, hardly, honestly, actually, sincerely
Space	place, position, direction	-
Time	temporal relations (fixed position in time, duration)	still
Other (not included in Scheibman's model)	all other instances of adverbs	by the way, also, anyway, although

Degree of epistemic commitment

The following factors are studied relative to the semantic content in main and complement clauses. Epistemic meaning can paradoxically display both certainty and uncertainty towards the proposition it modifies. For this, epistemic commitment is best illustrated by an epistemic scale. One end of the scale represents complete certainty and possibility and is therefore opposed to the other end with limited confidence in truth. The position that certain stance constructions take on the scale determines their degree of certainty. Therefore, examples typically accompanied by modals and other mitigating markers are coded *weak*, and constructions with such strengthening markers as adverbs, explanation marks or capital letters, are coded *strong*, as shown in (38) and (39) respectively. *Neutral* is the unmarked value of the factor, and the absence of any overt intensification or attenuation, like in example (40), classifies the construction as such.

(38) *I think, like Kaja, that she might try to come back again though...*
(39) *As much as I loved her London routine, I think this is one of the best performances I've seen Kondakova do with Ribbon in a LONG time!*
(40) *I have to say I think the judging was fair and consistent tonight, with one exception: Trikomiti's ribbon.*

Evaluation
Evaluation is judged upon the proposition's general tone, which might range from *negative* to *neutral* to *positive*, as shown in examples (41), (42) and (43) respectively.

(41) *It seems to me like chinese style: if you are at home you have to win whatever..SICK SICK place Baku!*
(42) *I think it's more popular in Russia than in France.*
(43) *I think she's undoubtedly one of the most talented gymnasts in recent years...*

Neutrality is treated as the unmarked value of the factor, and all instances that are unbiased towards positive or negative stance are treated as such. As example (41) shows, the evaluative value of mental predicate constructions does not necessarily need to be entailed by the mental predicate clause or the clause it modifies, but evaluative cues also need to be retrieved from larger contextual units.

Verification
Since epistemic and evidential modality show the writer's commitment to and reliability of the embedded proposition, the members of rsg.net are running the risk of being challenged in terms of truth and reliability. The risk is even higher when the topic concerns knowledge shared by members of the community, or in other words, when other members are able to verify the proposition at hand, as demonstrated in (44):

(44) *Unfortunately there doesn't seem to be videos of her performing that element on youtube.*

In (44), the utterance is coded *verifiable,* since the existence of the video(s) in question can be confirmed or falsified by anyone in the community.

Example (45), on the other hand, is purely an impression and opinion of a gymnast's routine, which exhibits the member's internal state of mind on a specific subject matter. As a result, the utterance is coded *non-verifiable*.

(45) *Based on what I've seen from Corbeil, Merkulova's routine seem juvenile.*

Emphasis

Coding for emphatic utterances follows two main criteria: orthographic and lexical. Emphatic orthography in bulletin board entries is marked by such features as capital letters, which in Internet language equal to raising one's voice, as well as the presence of exclamation mark(s). Lexical emphasis is typically displayed through the use of emotionally charged content words, such as *impossible, always, really, the best,* etc. In addition, a common technique used to emphasize the verbs *think* and *seem* is the insertion of the auxiliary *do* in front of the predicate, as in examples (46) and (47):

(46) *But I **do** think Naazmi's apparatus handling is better..*
(47) *Like someone said in the Videos section, it **does** seem to be missing a bit of "flow", but it's so good to see her not performing routines like a Junior.*

All other instances without any overt attempt to highlight parts of one's utterances are coded *non-emphatic*.

Argumentativity

Taking a stance towards gymnasts, judges and events in rhythmic gymnastics also entails a great deal of imposition from members with opposing views. Coming from all over the world, rsg.net users represent different views, experiences and attitudes towards what is happening in the sport. Unavoidably, this gives rise to arguments and verbal conflicts. Whether utterances are *argumentative* or *non-argumentative* is determined by the user's exposure to the rest of the community and his/her apparent readiness to be challenged and defied. Example (48) is an instance of an *argumentative* use of epistemic stance.

(48) *Also, what is all this issues with Spain placing higher than them? I think it's fair enough, Spain is willing to fight and is doing it right.*

Non-argumentative is the unmarked value of the factor, and utterances without any overt indication of contentious behavior are coded as such.

Hypotheticality
Hypotheticality is also treated as a two-fold phenomenon in the present coding schema. Therefore, the examples retrieved from rsg.net are either *hypothetical/irreal* or *non-hypothetical/real*, depending on the user's construal of a situation as certain or assumptive. In previous works dealing with hypotheticality (Ford and Thompson 1986; Sweetser 1990; Ford 1997), irreality has mostly been identified in conditional sentences, such as in example (49). In fact, these constructions are found to often emerge as politeness functions in face-to-face and interpersonal interaction.

(49) *[...] but I really do think that even if she hadn't had the knot, she would not have been "allowed" to take 1st...*

Another feature typically expressing *hypotheticality* is the presence of modal markers (*would*, *could,* etc.), exemplified in (50). Hypotheticality can also be considered to appear in comparative *it seems as if* constructions that Aijmer (2009) regards as expressions of irreality and uncertainy, as demonstrated in (51).

(50) *I think "wildest dreams"* ***may*** *be the term you are looking for...*
(51) ***It seems to me as if*** *she's accepted that this Olympics is not going to be fair, and just wants to do her best.*

Most instances found in the corpus are *hypothetical/real*, in which case the happening of an event is not called into question.

Subjectivity
The next factor makes a distinction between *subjectivity*, an individual's stance towards states of affairs, and *intersubjectivity*, the construal of stance as a shared phenomenon between speakers and writers. The relationship between the two concepts is thoroughly scrutinized in Nuyts (2001), Scheibman (2007), Verhagen (2005), etc. Although in the previous chapters I have argued for the dialogical nature of communication, which regards intersubjective relations between members of the community as the key factor in the construction of stance, the present factor treats

intersubjectivity in a more narrow sense. Therefore, utterances that simply express the user's personal view without any overt addressing of fellow members are coded *subjective*. *Intersubjectivity* is primarily identified in utterances where the addressee is either inherent in the bulletin board structure, which is made possible by the 'quote' button that allows messages to be directed at specific members. It can also be marked in messages with the second person singular *you,* as in (52), or in case of an explicit marking of usernames, as demonstrated in example (53).

(52) *I think **your** example is unstable, but it is important to remember that the definition of unstable doesn't include everything.*
(53) ***Guillermo**, I know how you feel I think...*

Involvement
In van der Auwera and Plungian's (1998) study, the authors have undertaken the classification of modality. As a result, they propose to divide the notion into epistemic and non-epistemic modality, which in turn makes a distinction between participant-internal and participant-external phenomena. Although the present study is only concerned with epistemic modality showing probability and possibility as opposed to ability and non-epistemic possibility, I have modified van der Auwera and Plungian's classification to fit the goals of the current work, and treated both phenomena as applicable to epistemic possibility. Consider the following examples:

(54) *yeah it seems the Court of Arbitration is the highest level one can get.*
(55) *I think perhaps her ball and ribbon routines will grow on me, but the other two - yawn.*

This factor is first and foremost concerned with the proposition being modified by *think* and *seem*. A proposition is coded *external* when the writer is addressing an issue beyond his/her reach, i.e. the writer can be perceived more as an onlooker rather than a participant, as seen in (50). However, when the proposition is *internal*, as in (51), the poster is personally engaged in the situation and acts as a participant of the event.

5.2.3 Extra-linguistic factors

Social rank
Since social hierarchy and power on an online bulletin board is one of the central objectives in the present study, coding for the social rank of the sender as well as the intended addressee is a necessary step. In Chapter 3.1.3, rsg.net members were divided into three hierarchical ranks depending on their status and activity on the board. Therefore, social rank of the sender is coded for three values: *R1*, *R2* and *R3*. Addressees share the same categorization with an inclusion of *P*, meaning *public*, which is applied if the intended addressee is not explicitly present and the post is targeted towards the whole board.

Dialect
The dialect of the sender is labeled either *AUS* (Australian English) or *UK* (British English).

Topic of discussion
The thematic groups in which rsg.net members can initiate threads and elaborate on the existing ones were presented in Chapter 3.1.2. The same groups determine the division of topics in the coding schema. The complete list of the 14 groups includes: News & Announcements, The Gym, Code of Points, Events & Results, Gymnasts & Groups, Leotards & Equipment, Music & Editing, Gymnastic Photos, Photo Announcements, Gymnastic Videos, Video Announcements, Old Photos & Videos, Off Topic and RSG.net.

Chapter 6
Results of the quantitative study

Following the coding of linguistic factors, Chapter 6 adopts two statistical techniques: Multiple Correspondence Analysis (6.1) and Binary Logistic Regression Analysis (6.2). Both analyses are conducted in the open-source software R. A summary of the results is presented in 6.3. The adoption of both techniques is made possible by the multifactorial usage-feature analysis of 729 stance constructions with *think* and *seem*, which were coded for 31 formal, semantic and extra-linguistic factors. The aim of the adoption of both exploratory and confirmatory statistical tools is to identify the factors that are believed to most efficiently contribute to the distinction between authoritative and tentative stance-taking across three social ranks.

6.1 Multiple Correspondence Analysis

The first statistical technique employed in the analysis of stance is Multiple Correspondence Analysis. In his article on exploring linguistic data with Correspondence Analysis, Glynn proposes the following definition of the methodology: "It is an exploratory technique that reveals frequency-based associations in corpus data" (in press). The analysis gives insight into the correlations between linguistic factors that would otherwise go unnoticed. However, it is also important to acknowledge the main limitation of the analysis. With Multiple Correspondence Analysis, linguists are able to identify patterns in the data but unable to establish their statistical significance. This means that an additional confirmatory analysis needs to be applied to the results obtained from the exploratory technique (see Section 6.2).

Table 11. Summary of factors and values in the Multiple Correspondence Analysis

Factor	**Values**
Social Rank	*Sender*, *Addressee*
Object Person	*P1*, *P2*, *P3Human*, *Non-Referential*, *NA*
Epistemic Type	*Opinion*, *Conviction*, *Prediction*, *Question*
Epistemic Commitment	*Weak*, *Neutral*, *Strong*
Argumentativity	*Non-Argumentative*, *Argumentative*
Pragmatic Intention	*Downtoning*, *Prominence*, *Complaint*, *Insult*, *Praise*

The Multiple Correspondence Analysis is a space reduction technique where complex sets of associations are reduced to a 2D map, in which the correlations between various data points are calculated and converted into relative distances (Glynn 2010b: 251). Values that are situated close to one another represent strong associations and values that appear far from each other represent weak associations. The size of the data point is also relative to the degree that value contributes to the structuring of the data. Larger points are indicative of values that are important to the structuring, whereas smaller points are less important. In the following sections, I identify the factors that form clear and distinct clusters and display strong associations with social rank.

In Chapter 5.2, stance constructions with *think* and *seem* were coded for 31 formal, semantic and extra-linguistic factors. Including all 31 factors in the Multiple Correspondence Analysis would result in an over-complicated model, where the chance of false associations is significantly increased (Glynn in press). Therefore, only six factors (with their values) are added to the exploratory analysis: Social Rank, Object Person, Epistemic Type, Epistemic Commitment, Argumentativity and Pragmatic Intention (see Table 11, above). The factors are chosen based on the observations made in the course of the coding process where certain patterns and co-occurrences arose. Therefore, the factors are expected to provide us with results that will indicate the authoritative and tentative nature of stance-taking. In addition, several tests were run with a variety of factors to obtain distinct patterns and the six factors presented above resulted in the clearest model. In line with the purpose of the present study, the most important factor added to the analysis is social rank, where both the rank of the sender as well as the addressee are included. They are incorporated in a way that displays the direction and relationship between the two parties in order to account for the dialogical nature of stance. For

instance, when a Rank 1 user addresses a Rank 2 user, it is represented as R1->R2, when a Rank 3 member approaches the whole board it is represented as R3->P, etc.

The results of the Multiple Correspondence Analysis with six factors are visualized in Figure 2 below. The R package used for the present analysis was {FactorMineR}. All factors on the biplot are studied relative to their association with social rank. Figure 2 shows that the y-axis of the biplot divides the ranks into two distinct groups. The space occupied on the right side only displays rank combinations with Rank 1 as senders, while the left side contains rank combinations with lower ranks as senders. The clear vertical distinction between Rank 1 users (i.e. moderators) on the one hand and Rank 2 and 3 users (i.e. hosts and casual senders) on the other shows that epistemic and evidential stance used by the highest rank in rsg.net contains distinct features that deserve to be explained. In the following sections, I distinguish between three clusters identified in Figure 2. Although some are more sound than others, careful conclusions can be drawn in all cases.

6.1.1 Cluster 1: R1->R2

The strongest association in the analysis can be found in the top-right quadrant of the biplot. The cluster is clearly detached from others, which means that in regards to the rank combination R1->R2, the factors are used in a specific way. As can be seen from the analysis, these factors include: Object Person with its value *P2*, Pragmatic Intention with *Insult* and *Argumentative* as the value for Argumentativity. Let us now look at each factor in isolation and try to explain their associations with the rank combination.

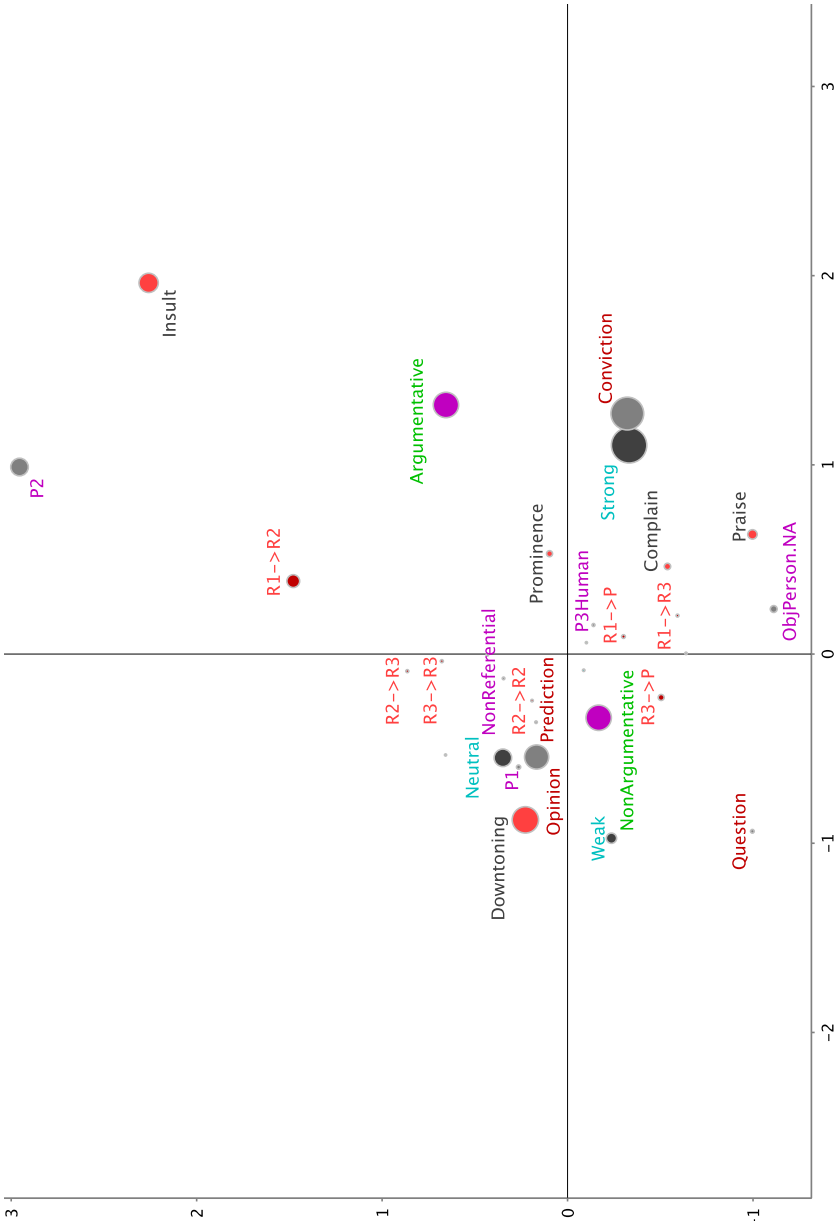

Figure 2. Multiple Correspondence Analysis: Social Ranks, Object Person, Epistemic Type, Epistemic Commitment, Pragmatic Intention, Argumentativity

The factor most distinctly drawn to the rank combination R1->R2 is Object Person and its value *P2*. The value is situated furthest from the center of the plot, which indicates its strong association with the rank. As established in Chapter 5.2.1, Object Person is identified in the subject position of the complement clause modified by the verb. *P2* therefore marks the second person singular and plural realizations of the factor, as in *I think you are wrong*. The fact that Rank 1 users address Rank 2 users with the pronoun *you* might imply that these ranks are most often involved in discussions and arguments where each party is explicitly addressed. The nature of these arguments is highlighted by two other factors the rank combination is associated with.

The first factor that sheds light on the content of the utterances produced by Rank 1 users is Argumentativit*y*. The value represented in the speech of Rank 1 users is *Argumentative*, which indicates the users' readiness to be involved in arguments and imposed on in later interaction. The second factor that illustrates the stance taken by Rank 1 users is Pragmatic Intention. This factor is exemplified by the value *Insult*. Of all values given in the factor, *Insult* is identified as the most severe confrontation between the three ranks. The fact that it is characteristic to the speech of Rank 1 users shows that besides being argumentative and upfront, moderators tend to discard basic politeness strategies and often confront their fellow members in a degrading manner. The fact that Rank 1 users tend to be argumentative with Rank 2 users shows that when addressing other experienced and active members of rsg.net, moderators constantly remind them about their privileged position on the board by using authoritative and reassuring stance constructions. Therefore, the co-occurrence of these values provides strong support to the adoption of assertive and dominating stance markers by Rank 1 users.

6.1.2 Cluster 2: R1->R3/R1->P

To continue with Rank 1 users as senders, we now take a look at the bottom-right corner of the analysis. As already mentioned, the y-axis running through the plot clearly distinguishes between moderators on the one hand and hosts with casual senders on the other. This suggests that the use of epistemic and evidential stance markers displays clear segregation between the ranks. While the previous cluster represented the rank combination R1->R2, i.e. a moderator addressing a host, this cluster

represents ranks R1->R3 and R1->P, where moderators now turn to the lowest rank, casual senders, as well as the rsg.net bulletin board as a whole. Although the bottom-right quadrant of the biplot does not provide as clear a cluster as the one above, a few observations can still be made. The factors associated with the two rank combinations are: Object Person, Pragmatic Intention, Epistemic Type and Epistemic Commitment.

Similar to the previous cluster, these rank combinations are also associated with values from the factor Pragmatic Intention. The values characterized by the two rank combinations in this case are *Praise* and *Complaint*. In fact, the functions of these factors are fully understood with regard to another factor in the cluster, namely Object Person. The interaction of these factors is explained by the interdependence of referents and descriptive features, and the necessary existence of an object being either praised or complained about.

Let us look at the association more closely. Object Person is represented in the cluster by two values: *N/A* and *P3Human*. *N/A*, as reported in Chapter 5.2.1, is only applied to *seem*-constructions with a copular phrase, in which case the clause succeeding the perception verb has no explicit subject (e.g. *She seems nice*). That construction clearly draws towards *Praise* and the two rank combinations in question. In addition, there is another value that can be identified near the two ranks. The value *P3Human* indicates the existence of third person singular or plural constructions with a human referent, which typically succeed *think*-constructions. However, the value can be found very close to the center and is therefore not considered to be distinctly associated with Rank 1 users as senders. Being close to the center implies that the factor is also common to other rank combinations found in the study. As a result, the association of *P3Human* with the rank combinations R1->R3 and R1->P is considered to be weak. However, the co-occurrence of these values might imply that the objects of praise by Rank 1 users are not fellow members, but rather gymnasts, coaches, judges or other people outside the virtual community.

Another factor associated with the rank combinations in question is the Pragmatic Intention *Complaint*. As can be assumed on the basis of the findings scrutinized above, the Pragmatic Intention can also be concerned with people outside rsg.net rather than other members of the bulletin board. Therefore, in addition to praise towards the 'heroes' of the sport, Rank 1 users are also likely to produce utterances with an illocutionary act of complaint.

Last but not least, the present cluster includes two values that did not occur in the previous cluster. These are *Strong* from Epistemic Commitment and *Conviction* from Krawczak and Glynn's Epistemic Type. Both values are relatively strongly associated with the rank combinations R1->R3 and R1->P. They are indications of the existence of authoritative and reassuring functions applied to epistemic and evidential stance. This means that when Rank 1 users turn to the public or to Rank 3 users, their language is characterized by strong and convincing constructions framed by *think* and *seem*.

6.1.3 Cluster 3: R3->P/R2->R2

The third cluster is represented on the left side of the plot and separated from other clusters by the y-axis. The rank combinations being scrutinized here have Rank 2 and 3 users as senders and the public and Rank 2 users as addressees. Although separated by the x-axis, the two rank combinations in question, R2->R2 and R3->P, seem to be associated with a number of shared values. A closer look at the values shared by the two rank combinations also reveals similar characteristic functions. The factors represented are: Object Person, Epistemic Type, Epistemic Commitment, Pragmatic Intention and Argumentativity.

All the factors associated with the rank combinations are clustered together in a more or less straightforward way. However, there is one factor that seems to pull away from the others. It means that the factor displays a strong association with its closest rank combination R3->P. The factor in question is Krawczak and Glynn's Epistemic Type with its value *Question*. As established in Chapter 5.2.2, the value is applied to epistemic and evidential stance constructions that function as questions. The association implies that when Rank 3 users turn to the public they are more likely to use questions than other ranks. The adoption of a question rather than an affirmative sentence structure indicates that casual senders aim to minimize the risk of committing themselves to their judgments. Instead, they turn to the board to look for confirmation from other members.

The rest of the values associated with the two rank combinations are mutually shared by both. Again, Object Person is represented in the cluster, but in this case with the values *P1* and *Non-Referential*. While the former deals with first person singular constructions where writers refer to themselves as the agents of the complement clause, the second value

includes such dummy subjects as *there*, *it*, etc. This might indicate that Rank 2 and 3 users are more likely to position themselves as the agents of the events being talked about, rather than pointing at others in discourse. In addition, the usage of such non-referential subjects as *there* and *it* reveals that these members tend to avoid addressing specific subjects and rather employ techniques where this can be avoided.

The other factors juxtaposed with Object Person give clear indications of the nature of messages produced by Rank 2 and 3 members. The clustering of such values as *Weak* and *Neutral* epistemic commitment, *Downtoning* and *Non-Argumentativity* shows that when addressing members from their own rank, Rank 2 and 3 users tend to use tentative epistemic and evidential constructions. In fact, these values can be opposed to those found in the messages of Rank 1 users above. Moreover, the close association of two values from Krawczak and Glynn's Epistemic Type, namely *Opinion* and *Prediction*, shows that these members are more likely to express 'pure opinion' and make careful predictions about the future than take a strong stance towards the events in rhythmic gymnastics.

It must be remembered that these results are not confirmed. Although Correspondence Analyses are powerful tools for identifying complex relations in the data, strong claims about language structure beyond the sample cannot be made. Moreover, for the most part, the social rank data points do not appear to be important to the contribution of the overall structure. This does not detract from the findings, but adds weight to the fact that caution must be taken in their interpretation. In order to determine which of the findings are significant and likely to be representative of the structures beyond the sample, Binary Logistic Regression Analysis is used.

6.2 Binary Logistic Regression Analysis

The next analysis applied to the factors of *think* and *seem* is Binary Logistic Regression Analysis. The R package used for this technique is {rms}. In contrast to the previous exploratory approach, Logistic Regression Analysis is a confirmatory technique that helps researchers test the statistical significance, effect size and predictive power of factors and values (Speelman in press). While Multiple Correspondence Analysis identifies patterns in data and highlights linguistic features that might otherwise go unnoticed, the present analysis verifies the results and is therefore a natural continuation to the exploratory technique.

The analysis produces three main scores. They are (1) the probability score of each factor, (2) their effect size, and (3) the predictive power of the model. First, the probability score or *p*-value provides the statistical significance of each factor by answering the following question: What is the probability that the feature in question will predict in a different manner if the study were to be repeated? In other words, it shows whether the results are merely a matter of chance or not. This score can be found in the last column of the table of coefficients, in which case any score beneath 0.05 is considered to be statistically significant, which is also indicated by the presence of one or more asterisks.

Secondly, the importance of each factor in structuring the model is determined by the estimates of the coefficients, which can be found in the second column of the table of coefficients. For an easier interpretation of the results, numbers higher or lower than +/- 1 are typically considered to be important. Moreover, in the Binary Logistic Regression Analysis, negative numbers predict for one outcome and positive numbers for the other.

Thirdly, the predictive power of the model shows how often it is possible to predict the outcome, taking into consideration the factors used in the analysis. The scores are provided at the bottom of each analysis. First, the Nagelkerke Pseudo R^2 is a calculation used to evaluate the goodness-of-fit of the logistic model. It should be noted that for the Logistic Regression Analysis, the true R^2 is not possible to be applied, and therefore a pseudo-model needs to be used. Any score above 0.3 is considered to be predictive (Lattin, Carrol, and Green 2003). Next, the concordance statistic or C-score is "an index of the correlation between predicted probability of expected response and actual response" (Glynn 2010: 259). A strong result is any value above 0.8 (Hanley and McNeil 1982). In addition, all factors were checked for multicollinearity, a factor that misleadingly magnifies standard errors (Speelman in press). Moreover, the highest Variance Inflation Factor is reported in all models. When the factor remains below 4, the consequence is not considered to be problematic. In all models, the number is lower than the critical figure, which means that the results are valid and unaffected by undesirable correlations between predictors.

The factors chosen for the confirmatory study are the following: Object Person, Pragmatic Intention, Argumentativity and Evaluation. The first three factors with their most distinct values are retrieved from the Multiple Correspondence Analysis and treated as factors that most

successfully participated in the structuring of the model. In addition, several tests were run with factors not included in the exploratory analysis to gain better insight into other factors affecting social rank. Although not represented in the previous technique, the factor Evaluation still seems to be an important indicator of social rank in the messages of rsg.net members. In the following sections, I present all four factors separately and attempt to explain the results obtained through statistical modeling.

6.2.1 Model 1: Object Person

The factor Object Person in the Multiple Correspondence Analysis showed that Rank 1 members were more likely to address Rank 2 members by using the second person pronoun *you*, while addressing Rank 3 users and the public showed associations with *N/A* and *P3Human*. When Rank 2 and 3 users address their interlocutors they tend use the values *P1* and *Non-Referential*. Some of these results are now tested using the Binary Logistic Regression Analysis to either confirm or disprove the patterns identified above.

First, I introduce two terms that are important for the interpretation of confirmatory results: outcome and predictor. Outcome is a binary set of values that is predicted by the factors included in the analysis. These factors are called predictors. In the first model, outcome is Object Person, and more specifically, its values *P1* (first person singular) and *P2* (second person singular or plural). The two values are chosen for their associations with the highest and the lowest social rank on the exploratory biplot. The predictors are listed in the first column of the table of coefficients. The predictors in this study are the rank combinations. Therefore, the model attempts to show which rank combinations predict the use of first person subjects in the complement clause and which combinations predict second person subjects.

Table 12 shows that three rank combinations are marked as statistically significant and one as borderline significant (indicated by asterisks). The two strongest predictors are R1->R2 and R3->R3 as both have *p*-values that are below 0.005, which is a strong result. In establishing how important the predictor is, we need to look at the second column of the table, or the estimates of the coefficients. As already mentioned, negative numbers predict for one outcome and positive numbers for the other, and numbers higher or lower than +/- 1 are generally considered to be

important. In the present case, all numbers with a negative value predict *P1* (first in the alphabet) and all positive values predict *P2* (next in the alphabet). Therefore, it is possible to conclude that all significant predictors are predicting *P2* or the adoption of the second person singular or plural subject in the complement clause. Also, all significant rank combinations have estimates above 1 and therefore act as important factors in predicting Object Person.

Table 12. Binary Logistic Regression Analysis: Object Person

```
Object Person: P1 (72 examples) and P2 (26 examples)
Predictor Factor ~ Rank Difference

Coefficients:
            Estimate    Std. Error   z value    Pr(>|z|)
R3->R2        0.6931        1.4719     0.471     0.63771
R1->P        -0.2451        1.4486    -0.169     0.86563
R1->R2        3.2958        1.1547     2.854     0.00431 **
R2->P         1.9095        1.2304     1.552     0.12066
R2->R2        2.4849        1.3743     1.808     0.07060 .
R2->R3        2.8904        1.4337     2.016     0.04380 *
R3->R3        3.7377        1.2376     3.020     0.00253 **

    Null deviance: 113.393  on 97  degrees of freedom
Residual deviance:  80.225  on 90  degrees of freedom

Nagelkerke Pseudo R2: 0.419 (randomized n= 1000: 0.3582)
Concordance Statistic: 0.845 (bootstrapped n= 1000 0.784)
Highest variance inflation factor: 3.427834 (R1->R2)
```

The most significant rank combinations predicting the second person singular or plural *you* are ranks R1->R2 and R3->R3. The first rank combination also established a strong association in the exploratory analysis, where the factors formed a clearly visible cluster. The confirmatory technique reveals that the association is not merely a chance and with a 3.2958 estimate, the combination is a strong predictor of *P2*. The combination R3->R3, however, did not become apparent in the previous analysis, but proves to be an important predictor with an estimate of 3.7377. Looking at the other rank combinations with statistical significance, a pattern seems to emerge. All rank combinations that *P2* is predicted by are combinations where the direction of the sender and addressee is either equal or 'pointing downwards'. This means that when moderators, hosts or casual users act as senders, they are more likely to use the second person singular or plural subject to refer to another member

from an equal or lower rank. Rank combinations where the relationship is directed upwards are not even statistically significant. This might refer to the fact that rsg.net members are more comfortable to directly address their interlocutors belonging to the same hierarchical level or the level above them.

As indicated above, confirmatory techniques do not only test statistical significance but also measure the predictive power of the model. These scores demonstrate how accurately it is possible to predict the outcome with the features at hand. It therefore shows how well we can predict the factor as either A or B based on social rank. The scores are presented at the bottom of each model. As established above, any score above 0.3 in the Nagelkerke Pseudo R^2 is a strong result, which means that the score 0.419 confirms the goodness-of-fit of the model. In addition, the C-score of the model is above 0.8 (0.845), which shows that the model has strong predictive strength.

6.2.2 Model 2: Pragmatic Intention

Model 2 deals with the semantic factor Pragmatic Intention. As indicated in Chapter 5.2.2, the factor consists of a variety of levels, ranging from humor to insult. In the Multiple Correspondence Analysis, Pragmatic Intention formed strong associations with a number of rank combinations, the most obvious of them being the clustering of R1->R2 and *Insult*. For the sake of comparison, the two values chosen for the present analysis are *Insult* and a value that can be considered to be its opposite, *Downtoning*. The latter value in the exploratory analysis was associated with rank combinations where senders were Rank 2 and 3 users. However, as the cluster was formed close to the center, no firm conclusions could be drawn.

As can be seen in Table 13, the only statistically significant rank combination predicting Pragmatic Intention is R1->R2 (*p*-value=0.00784). The coefficient estimate in the second column shows that with an estimate of 1.44238, the predictor is strong and therefore an important predictor of the pragmatic value. Since the number is positive, it predicts *Insult*, which confirms the strong association seen in the Multiple Correspondence Analysis. With the adoption of the confirmatory technique, it becomes clear that the strong correlation seen in the exploratory analysis is not a matter of chance and that it is being verified using complicated statistical analyses. Therefore, we can confidently say that when moderators turn to hosts, their

language is characterized by aggression and insult towards their interlocutor. Moreover, the weak association between *Downtoning* and such rank combinations as R3->P and R2->R2 is verified by the confirmatory technique.

Table 13. Binary Logistic Regression Analysis: Pragmatic Intention

```
Pragmatic Intention: Downtoning (274 examples),
                     Insult (31 examples)
Predictor Factor ~ Rank Difference

Coefficients:
           Estimate  Std. Error   z value    Pr(>|z|)
R1->R2      1.44238     0.54250     2.659     0.00784 **
R1->R3      0.09531     1.12064     0.085     0.93222
R2->P       0.45199     0.62001     0.729     0.46601
R2->R2     -1.21302     1.08759    -1.115     0.26471
R2->R3      0.52609     0.85602     0.615     0.53883
R3->P      -0.69315     0.71016    -0.976     0.32904
R3->R2      0.78846     0.86939     0.907     0.36446

    Null deviance: 200.49  on 304  degrees of freedom
Residual deviance: 184.10  on 297  degrees of freedom

Nagelkerke Pseudo R2: 0.130 (randomized n= 1000: 0.0833)
Concordance Statistic: 0.725 (bootstrapped n= 1000 0.66395)
Highest variance inflation factor: 1.538899 (R1->R2)
```

The Nagelkerke R^2 index at the bottom of the model reveals that the predictive power of the model is rather weak. The score 0.130 is well below the satisfactory level. However, as the C-score is considered to be a more important and reliable indication of predictive power, this should also be taken into consideration. However, the score 0.725 is relatively low compared to the previous model, which confirms the rather weak predictive power of social rank in determining Pragmatic Intention.

6.2.3 Model 3: Argumentativity

The next factor to be scrutinized is the semantic factor Argumentativity (Table 14). Argumentativity is also familiar from the Multiple Correspondence Analysis, where both of its values, *Argumentative* and *Non-Argumentative*, were represented. In the exploratory analysis, the former value was strongly associated with the rank combination R1->R2.

Alongside *Insult*, the factor established the strong and aggressive use of epistemic stance by the rank combination. *Non-Argumentative*, however, was not considered to be a distinct factor, but was found in close proximity to Rank 2 and 3 users as senders.

Table 14. Binary Logistic Regression Analysis: Argumentativity

```
Argumentativity: Argumentative (144 examples)
                 Non-Argumentative (572 examples)
Predictor Factor ~ Rank Difference

Coefficients:

          Estimate  Std. Error   z value    Pr(>|z|)
R2->R3     0.15301     0.53013     0.289    0.772871
R3->R3     0.30608     0.45113     0.678    0.497470
R2->R2     0.16793     0.34402     0.488    0.625449
R1->R2    -0.65590     0.29932    -2.191    0.028432 *
R1->R3     0.16793     0.45640     0.368    0.712909
R2->P      0.09894     0.30380     0.326    0.744671
R3->P      1.18958     0.32634     3.645    0.000267 ***
R3->R2     0.42744     0.51808     0.825    0.409340

    Null deviance: 718.79  on 715  degrees of freedom
Residual deviance: 689.59  on 707  degrees of freedom

Nagelkerke Pseudo R^2: 0.063 (Bootstrapped: 0.0290)
Concordance Statistic: 0.629 (Bootstrapped 0.5964)
Highest variance inflation factor: 1.183439 (R2->R2)
```

The confirmatory analysis of Argumentativity detects two statistically significant rank combinations. The combination R1->R2 that previously formed a clear cluster with the *Argumentative* value of the factor also turns out to be significant in the present analysis (p-value=0.028432). However, the rank combination R3->P for which only careful conclusions were drawn in the exploratory analysis, is followed by three asterisks, i.e. it is a highly significant predictor of Argumentativity (p-value<0.0003). Looking at the coefficient estimates, the outcome being predicted by these rank combinations are in line with the results obtained in Section 6.1. The first significant combination, R1->R2, is a predictor of *Argumentativity* and R3->P a predictor of *Non-Argumentativity*. However, as the estimate for R1->R2 is below 1, the predictor is not considered to be very important. On the other hand, R3->P exhibits a coefficient estimate higher than 1 and is therefore an important predictor of the outcome.

The results show that Rank 1 senders use argumentative epistemic stance constructions with members who are also experienced and active members of the board, but who nevertheless have not been promoted to the highest position. On the other hand, Rank 3 users address the board with tentative and careful epistemic constructions and therefore minimize the risk of being imposed on by more experienced users of the board.

Similar to the previous model, both scores evaluating the predictive power of the sample are beneath the satisfactory level, 0.063 and 0.629 respectively, which implies that Argumentativity is not accurately predicted by the three social ranks.

6.2.4 Model 4: Evaluation

The last factor to be tested using the Binary Logistic Regression Analysis is Evaluation. This factor is not found in the Multiple Correspondence Analysis, since it did not contribute to the distinct clustering of factors. However, after running a series of confirmatory tests on a number of factors annotated in Chapter 5.2 but not represented in the exploratory analysis, Evaluation turned out to give the most interesting results (Table 15, below).

First, statistical significance of the factor relative to social rank is promising. Altogether four rank combinations have *p*-values lower than 0.05. The predictor with the lowest *p*-value (0.000946) is the rank combination R3->P. This is followed by ranks R2->P and R3->R2 with *p*-values lower than 0.05, and R3->R3 with borderline significance. What unites all these predictors is their correlation with one outcome only. In Chapter 5.2.2, Evaluation is listed as a three-fold factor with the following values: *Negative*, *Neutral* and *Positive*. As the two opposing values, *Negative* and *Positive*, were expected to give better insights into the distinction between the three ranks, these values were added to the confirmatory analysis. Therefore, since all estimates are positive numbers, only the positive value of the factor is being predicted. The most important predictor according to the table of coefficients is R3->P with a value of 1.6007, closely followed by the most statistically significant rank combination R3->P. This confirms their importance in predicting the positive outcome of the factor.

Table 15. Binary Logistic Regression Analysis: Evaluation

```
Evaluation: Negative (180 examples)
            Positive (188 examples)
Predictor Factor ~ Rank Difference

Coefficients:
         Estimate Std. Error   z value     Pr(>|z|)
R2->R2     0.3967     0.4887     0.812     0.416874
R1->R2    -0.1141     0.4606    -0.248     0.804352
R1->R3     0.5509     0.5513     0.999     0.317636
R2->P      0.8650     0.3823     2.263     0.023666 *
R2->R3     0.9076     0.6790     1.337     0.181342
R3->P      1.1246     0.3402     3.306     0.000946 ***
R3->R2     1.6007     0.6603     2.424     0.015345 *
R3->R3     1.0411     0.5735     1.815     0.069464 .

    Null deviance: 426.83  on 312  degrees of freedom
Residual deviance: 406.79  on 304  degrees of freedom

Nagelkerke Pseudo R2: 0.083  (Bootstrapped: 0.0253)
Concordance Statistic: 0.641  (Bootstrapped 0.6015)
Highest variance inflation factor: 1.576021 (R3->P)
```

A pattern that emerges after a brief investigation of the predictors indicates that positive stance is only used by rank combinations where senders are rsg.net members from Ranks 2 and 3. No instances of positive stance can be found from combinations with moderators as senders. This demonstrates that as opposed to the privileged rank of the board, users who have not institutionalized their role in the community tend to be more amiable and sympathetic towards their fellow members.

Again, the predictive power of the model, in which social rank is predicted by Evaluation, is weak. Both numbers, the Nagelkerke Pseudo R^2 and the C-score are below 0.3 and 0.8, which implies the low accuracy of the model.

6.3 Summary

In Section 6.1, an exploratory technique to statistics was adopted to identify the patterns in the data and show the associations between epistemic stance and social rank. Three distinct clusters could be identified. The cluster that formed the most distinct associations was Cluster 1, in which the rank combination R1->R2 was strongly associated with the second person

singular or plural subject in the complement clause, and insulting and argumentative stance constructions. This demonstrates that when addressing other experienced members of the board, moderators tend to display aggressiveness and use epistemic and evidential stance verbs to frame strong and authoritative propositions, as exemplified in (1).

(1) *Youandi you can stop NOW with this wingy tune cause I don;t think there's such a HUGE chaos of spamming topics left around.*

This message is posted by the Rank 1 user *Storm*, whose response to the Rank 2 user *Youandi* carries all the markers identified in the cluster: the second person singular *you* with an argumentative and insulting tone. Although *Youandi* is an experienced and active member of the board, who has produced more than 10,000 messages since joining rsg.net, the moderator does not seem to show much concern towards his/her face wants.

The Binary Logistic Regression Analysis adopted in 6.2 supports and also brings into question some of the results obtained through exploratory modeling. For instance, the weak association between the rank combination R3->P and the factor *Non-Argumentative* in the exploratory analysis was found to be statistically significant in the confirmatory technique. This does not mean that the interaction of features identified in the Multiple Correspondence Analysis is invalid, but rather draws attention to the relatively small size of the sample. In the exploratory technique, small samples are immensely sensitive and prone to form patterns, which can be falsely interpreted. The confirmatory analysis is more resistant towards reaching statistical significance with a small number of values, which explains the differences encountered in the two analyses. However, the fact that significance was achieved in all models shows that these factors of epistemic and evidential stance are strongly dependent on social hierarchy and power in the community.

Taking into consideration both the scores of statistical significance and predictive power, the most important outcome of social rank seems to be Object Person with its values *P1* and *P2*. All significant predictors or rank combinations were found to predict only one of these outcomes, namely second person singular or plural. The statistical significance of the rank combination R1->R2 confirms the results obtained from the Multiple Correspondence Analysis, in which the two factors were undeniably associated. Another significant predictor is R2->R3 where the direction

again points downwards towards a lower rank. This implies that rsg.net members are more comfortable in addressing fellow members with the personal pronoun *you* from ranks lower than their own. In addition to this, another pattern seems to emerge from the results. This concerns other statistically significant predictors such as R3->R3 and R2->R2 that also predict the outcome *you*. These factors did not form a distinct cluster in the exploratory analysis, but show significance in the Binary Logistic Regression Analysis. The outcome can be explained by the fact that when hosts and casual senders turn to members from their own rank, no apparent power conflicts are present. As both members belong to the same social rank, they are perceived as equal and impartial, which means that in addition to displaying dominance, the second person *you* in the complement clause can also be used as a marker of solidarity and comradeship. Consider the following example by the Rank 3 user *Cameron* in addressing another casual sender *alex_abt*:

(2) *How lucky Alex - I think you're the only person on here to get the AA final!*

The remaining three factors, Pragmatic Intention, Argumentativity and Evaluation, are also statistically significant and establish a strong contrast in the use of epistemic stance by moderators, hosts and casual senders. As the predictive power of the three models is below the critical level (0.3 in Nagelkerke R^2 and 0.8 in the C-score), the results need to be treated with care. However, it should be borne in mind that the aim of the present study is not to predict the various formal and semantic factors framed by *think* and *seem* by solely relying on social rank. Instead, the goal of the quantitative study is to identify the factors that are distinctly associated with social rank in bulletin board messages. These results are established by the scores of statistical significance, which were found to be important in all four models.

In summary, the factors that indicate social rank in the quantitative part of the study are: Object Person, Pragmatic Intention, Argumentativity and Evaluation. While a number of patterns emerged from the Multiple Correspondence Analysis, it is the Binary Logistic Regression Analysis that eventually determines the statistical significance of the outcome. Since statistical significance is influenced by sample size and the sample size of constructions with *think* and *seem* is too low from which to draw firm conclusions, the emergence of any significance shows the importance of

these factors on social rank. The results also show clear distinction in the use of epistemic stance by moderators, hosts and casual senders, where moderators are more likely to address members from lower ranks with dominance and aggression, while hosts and casual senders use epistemic and evidential constructions in more tentative contexts.

Chapter 7
Conclusion

The aim of this thesis was to explain the socio-cognitive functions of epistemic and evidential stance verbs on an online bulletin board. The stance markers were studied from the perspective of social power and hierarchy, aspects that are believed to have an influence on the construction of stance in the virtual community. Two different analytical tools were used to investigate the influence of social rank on bulletin board messages: the qualitative dialogical discourse analysis and the quantitative multifactorial usage-feature analysis. The adoption of these distinct but complementary cognitive models was expected to give better understanding of epistemic and evidential stance-taking in CMC.

The data for the present study were retrieved from the bulletin board rsg.net. The board is specialized on rhythmic gymnastics and attracts gymnastics fans from different parts of the world. Social hierarchy on the board is operationalized based on two criteria, status and activity, and rsg.net members were divided into three social ranks: moderators, hosts and casual senders. Moderators have a privileged role in the community and were assigned the highest position because of their status on the board. Hosts and casual senders were identified based on their activity in rsg.net discussions, in which case the members who had contributed more than 700 messages to the board were categorized as hosts, and members who had contributed fewer than 700 messages were assigned the role of casual senders.

The two analytical methods adopted in the present study aim to answer one overarching research question and two questions applied to each of the methods. The former sounds as follows:

— Do Internet bulletin board members of a higher rank use epistemic and evidential verbs in a more assertive and authoritative manner than users of a lower rank?

Answers to this question are presented using two analytical methods (see below). First, the qualitative approach attempts to answer the following research question:

— How do members from three different ranks attenuate or reinforce the strength of their propositions relative to their conversational co-participants?

The method adopted in the qualitative study, dialogical discourse analysis, assumes that language is a dialogically constructed phenomenon and that stance is co-constructed between members of the communicative act. By looking at all epistemic and evidential stance verbs in two controversial threads retrieved from rsg.net, two main conclusions were drawn: (1) rsg.net members align their stance constructions with fellow members, and (2) the representation of stance constructions across three different ranks shows negligible differences.

The first result confirms the presence of epistemic alignment between two or more participants. The analysis shows that rsg.net members build their epistemic and evidential stance constructions on previously expressed stance markers. The stance markers typically modify the same propositions across speakers, which means that when a writer positions himself/herself relative to a certain person, event or state, successive writers are likely to take a stance towards the same matter.

The second result demonstrates that qualitative approaches to stance-taking provide us with interesting and novel representations of stance in context, but only cautious generalizations can be made on its authoritative or tentative nature. The most distinct epistemic use can be detected in the messages of casual senders. These users were most often found to be pointing forward in discourse by first acknowledging the dominant views present on the board, which were then followed by the users' own tentative views on the matter. The members, therefore, minimize the risk of being challenged later in the discussion. This is in line with the general assumptions behind dialogicality, which does not only account for previous turns but also possible new turns that might emerge from interaction.

The two moderators present in the messages of the two discussion threads did not produce any epistemic or evidential stance verbs. At the same time, their messages were considered to be authoritative and dominating and were marked by different linguistic strategies. Therefore, it was concluded that more than the rank of the conversational participant, the degree of epistemic and evidential stance is influenced by the communicative situation in which rsg.net members found themselves. This means that the use of epistemic stance by a casual sender might vary from tentative to authoritative depending on the nature of previous confrontations. Therefore, due to the relatively small sample size and a high variety of stance used by the three ranks, no definite conclusions can be drawn from the analysis. For a more thorough overview of the nature of epistemic and evidential stance verbs, more examples need to be extracted from the bulletin board.

Another method, the multifactorial usage-feature analysis, was adopted to answer the following research question:

— What formal, semantic and extra-linguistic factors framed by *think* and *seem* indicate social hierarchy between the three ranks?

The primary aim of the analysis is to account for the interaction of all usage features. The interaction between morphology, syntax, lexis and social factors are all believed to be of importance in the construction of epistemic and evidential stance verbs on the board. In the quantitative part of the study, two frequently occurring stance constructions, *think* and *seem*, were selected and investigated relative to 31 formal, semantic and extra-linguistic factors. The manual coding of 729 examples resulted in a thorough usage profile of the two stance constructions. The factors that were believed to describe social hierarchy in the community most effectively were tested using two statistical techniques, Multiple Correspondence Analysis and Binary Logistic Regression Analysis.

The most distinct cluster in the Multiple Correspondence Analysis was formed around the rank combination where moderators address hosts with the following factors: Object Person, Pragmatic Intention and Argumentativity. The values of these factors showed that when moderators address users from the lower rank, their messages are characterized by insulting, argumentative and direct stance-taking. The other ranks, hosts and casual senders, were both grouped together with values that display weak argumentativity, neutral and weak epistemic commitment, and ‘pure

opinion' about the people and events in rhythmic gymnastics. To confirm the results obtained through exploratory modeling, three factors (Object Person, Pragmatic Intention and Argumentativity) that most clearly contributed to the structuring of the exploratory model and one factor (Evaluation) that was expected to differentiate between the three social ranks were included in the Binary Logistic Regression Analysis. The results showed that all four factors indicate the authoritative and reassuring use of epistemic and evidential stance by the highest rank. It was therefore concluded that when moderators address lower ranks, they primarily exhibit an argumentative and insulting style accompanied by the adoption of the second person singular or plural *you* to show dominance and power. On the other hand, when hosts and casual senders use epistemic and evidential stance, their language is non-argumentative and contains positive evaluations about their fellow members and people outside the Internet community. In addition, when addressing members from their own rank or the rank below them, hosts and casual senders also employ the pronoun *you*, in which case the pronoun is not used to establish social hierarchy and dominance but to rather show solidarity and equality.

The two methods described above are used to answer the overarching research question, which examines the authoritative and tentative use of epistemic and evidential verbs relative to social rank. It was found that the highest rank of the board, or moderators, tends to use language that can be categorized as strong and authoritative. Moderators not only use epistemic stance to exhibit power, but they also employ a considerable amount of stance adverbs and emotionally charged content words for this purpose. The difference between hosts and casual senders is less clear. Although hosts are rather experienced and active members of the board whose long involvement in the conversations of rsg.net has earned them an important place in the community, their use of epistemic stance verbs is not very different from the casual senders or 'newbies' of the board. The language displayed by these ranks is characterized by tentative and non-argumentative realizations of stance, where acknowledgements about the general viewpoints of the board are made before expressing their own tentative views on the matter.

The results of the two different methodological approaches complement each other in numerous ways. First, the qualitative study accounts for the behavior of epistemic and evidential stance verbs in their dialogical context, while the quantitative study identifies their usage features. Secondly, the qualitative study investigates a small number of

epistemic and evidential verbs in detail, while the quantitative study creates usage profiles for hundreds of examples, which are later applied to statistical modeling. Following Hunston's (2007) suggestion to apply both qualitative and quantitative methods to the investigation of stance-taking, the present work has managed to study stance from two perspectives and reached a general conclusion: expressions of epistemic and evidential stance verbs are important linguistic tools in highlighting social hierarchy and power in computer-mediated communication.

For future studies, it is important to improve the statistical techniques adopted in this study and employ an Ordinal Logistic Regression Analysis with Mixed Effects. This confirmatory method enables to predict all three levels of social rank. This means that instead of using social rank to predict only two usage features, it is possible to predict various combinations of social rank by a number of formal, semantic and extra-linguistic factors. Moreover, the models presented in this study did not control for possible influences of individual variation, dialect variation, or differences between the two constructions. To a certain extent, this shortcoming can be overcome by treating these variables as random, permitting predictions that factor out their potential influence. Finally, the investigation of epistemic and evidential verbs should not be the only approach to the subjective and impersonal language displayed in the virtual community, but rather serve as the first step in uncovering its undeniable abundance of stance-taking.

Appendix

Extract 1: Over-scoring

BRIVIDO (3)[8]: When she was a junior, everybody loved her. Now, everybody says she's a clown. Can you explain from where comes all this hatred?

ALEXANDRAFAN (2): There will always be immature people who hate a gymnast because of their style, and **I know** a lot of people don't like her style. She's more bright and bouncy as opposed to graceful and mature. Not everyone likes that. **I remember** one person saying a while back that if she wasn't overscored, everyone would be saying how much potential she had. But yes, mostly people hate her because she's overscored. Which **I think** is stupid; hate the judges or hate the coach, not the gymnast. She's just doing what she's told. Anyway, she's my all-time favourite gymnast. I can't stand gymnasts like Miteva or Kanaeva. I like gymnasts with charisma.

BRIVIDO (3): Ah what a pleasure to read you. You have a such intelligent reasoning. Like you, I don't like gymnasts like Miteva, Kanaeva, Staniouta, who bother more than anything, as you said,Merkulova have a lot of charisma (it's why she's often compared to Kabaeva **I think**).

ALEKSANDRAFAN (2): The funny thing is: **I don't see** Kabaeva having charisma. Aside from their similar faces, **I don't see** many similarities between Kabaeva and Sasha. I would rather watch something exciting and fun, with maybe some mistakes, than something perfectly clean but completely dead. I also love Maksimenko, especially her marvellous hoop from last year.

[8] Rsg.net username and social rank.

KALINKA (2):	+1. Except for their happy style, they are not similiar. **I neither find** Kabaeva charismatic (in 2001 and later), her routines had anything interesting, IMO. Whereas **I find** Merkulova very talented and charismatic, and she improves very fast (there is no comparison among her 2011 and her 2012 routines, but also among her routines in the start of the season and the routines she performed in Nizhny Novgorod). If she keeps improving so much, **I think** her 29 scores would be deserved. But, Brivido, it's not right that when she was a junior everyone loved her, because even in her very old videos people commented she was a clone of Kabaeva, that she had a fake smile and something similiar...
BRIVIDO (3):	It was really interesting to me to see different visions, and arguments, thanks to everyone I just wanted to say that when I wrote"hate" in the title it was because I saw so many insults against here on youtube! It made me so angry, because she's a human being, she's just 16, she worked very hard to be where she is now, and some people presume to blame her gratis like a bad film... It's why I wanted listen you to understand better . So, to resume why a lot of people don't like her : * Merkulova is the little dog of Viner; **I think** that's right, Viner has always had a favourite gymnast, but we can't do anything, unfortunatly... * She's overscored... **I don't know**, maybe I'm always optimist, but I can't imagine judges being corrupted :/ **I think** she's not as clean as Charkashyna, or Miteva, but her difficulty level is higher (because the cop is mainly made on the russian gymnasts...) * her smile is fake : of course!! like every gymnast ! What gymnast would be able to smile "real" during a so difficult effort?? **I also think** it contributes to

		the Merkulova's charisma, her youngness, her innocence, her immaturity, but youngness & innocence are not eternal.. Look at Kabaeva after 2001.

Extract 2: Maybe you are blind

YBALKA_ (2):	HAHAHAHAHAHAHAHAHA I would love to live in your bubble!
BRIVIDO (3):	^^ Come, it's rather cool! No... really, **I don't think** judges are corrupted! When people are angry because their fav gymnast doesn't win, it's normal. But don't call it injustice. Howerer.. **I think** judges are always guided by their unconscious, like every human being . It's natural.
KALINKA (2):	Really you can't imagine judges being corrupted? They definitely are. Also Sasha knows she doesn't deserve so high scores. Her D level isn't higher than Charkashyna's and Miteva's ones. **I'm quite sure** Liubov and Silviya start from 10, so Aleksandra can maximum start from the same value...
STORM (1):	!!!!!!?????? Maybe YOU are blind because YOUR favourite does win, then is easy for you to think "oh judges of course are fair!" Merkulova RUN, did a real marathon, to get the ball in her EC routine and she got 9.6 as execution. Now, you tell me. is this fairness??? Kanaeva for ages did the chest spin with both hands as help against the rules written in the COP, and she always got it counted. Mitroz and Weber are in the top 10, group A with super EASY routines, placing in front of Staniouta and Ritardinova etc- Is this fair judging? ALL the senior russian turn on almost flat foot most of the times, and they had the Cop changed to suit them. You yourself wrote that the CoP is written for Russian gymnasts: doesn't this

sound as the MOST UNFAIR possible thing? is this fine to you? Let's not be ridiculous please. Judges are UNFAIR big times. They are a shameful lot that plays as marionette with no conscience.

Extract 3: It's never too late

OLYMPIANRG (3): Ok, **I know** it sounds old to start RG at age 19, but I am really determined and prepared to work hard. Over the past few years I've become seriously enamored with this sport. **I think** it's one of the most beautiful sports there is. Please don't be confused by the 'Oympian' in the title, **I know** aiming for gold might not be possible, but just getting in the Olympics would already be heaven for me. I'm a perfectionist at everything that I do. I haven't followed any RG courses yet, as I wanted to focus on school first and my parents wouldn't let me. They said I still 'd have plenty of time after high school. I can already do all the splits with a slight oversplit on the left. I weigh 49kg and am 1.75 meter, which is the same as Anna Bessonova . I can also do a bridge and get my hands to about 15cm of my feet. I got a bit scared when I see people retiring at 23 years old . Why do they do that? Can't they still get better at RG? I would never give up. I'm from Belgium, so I'm looking for the best place to practice and a good coach. Could you guys please help me. **I know** it might take years and years to get to my goal, but I'm not a quitter and am prepared to work very hard. I'm already stretching about 5 hours a day.

SASTA33 (3): **I know** there's going to be some people who will say that 19 is a bit old to start, but I am not one of those people. **I think** that if it's an activity that you truly love, then you should go

for it no matter what. No one knows what the future brings, and **I think** that although other gymnasts start earlier, that doesn't mean that you shouldn't at least make a go for it. You just never know what could happen, and **I think** that's the beauty of sports. It's unpredictable. Who knows? Maybe you will find yourself to be a natural talent at RG, and will progress so rapidly that you'll surprise even yourself. You never know until you try. Sports, especially RG, should never just be available for little kids and teens. It should stretch through the adult years because a certain maturity and expressiveness comes with age. Look at Anna Bessonova. Look at Almudena Cid. Look at the artistic gymnast Oksana Chusovitina, who at age 34 is still going strong and wants to try for her sixth Olympic Games. Look at French gymnast Benoit Caranobe. He once said that "You are a gymnast all your life." I take those words to be absolutely true and inspiring. I am 25 now, and I still do RG, which I didn't start until age 13. I have improved more in my late teens and twenties than I ever did as a young teen, and I'm totally serious about this. **I understand** how to stretch correctly now, and how to work with the apparatus in more challenging ways, than I ever did as a young teen. It all makes more sense now, it really does. It's like now, more than ever, I "get it." **I think** those gymnasts who retire early retire because they had spent their whole childhood and adolescence in the gym, and then they got burned-out. Maybe tired, overworked, needed a break...who knows? Maybe they were placed into the sport by their parents before they, as people, really knew what they actually enjoyed doing. But you're in a unique position

because, as a young adult, you have the chance to pursue something that you're PASSIONATE about and can actually understand through the eyes of a young adult. It's going to really mean something to you because you know not to take it for granted. You found it "later" in life (sports-wise only, I mean!), and so you realize the preciousness of it. Age should play no factor in what you love to do. Go for it and see where it takes you. It sounds like you're a very hard worker already, so see what happens. Start with small goals and work your way up...try for smaller local competitions or shows, and then see how you do. Then aim a little higher and give that a try. Little steps. See where this takes you, whether that's the Olympics or simply fulfilling your goal of becoming a great gymnast. Either way, that's an amazing thing. It's so, so, so much better to give something a try and see what happens...then to never try at all and wonder what could have been.

KALINKA (2): **I suggest** you to start RG! It's never late for practising the sport you love! But (I don't want to be cruel) **I don't think** you will be able to attend very important competitions like Europeans, Worlds and Olympics... Gymnasts who partecipate to that competitions do RG since when they are child. For example, Russian, Belarusian, Bulgarian and Ukrainian gymnasts start RG at 3, 4 or 5 maximun. But **I know** that the Italian gymnast Susanna Marchesi started to do artistic gymnastics at 8 and at 11 she started rhythmic gymnastics. Anyway, 19, in my opinion, is too late for going to Olympics. It's never late to have fun, but not to have that high ambitions. Even if you are quite flexible and thin, you don't have the handling of the apparatus and

	the right body technique... But it should not be a problem if you love RG. If you love a sport, it doesn't matter what goals you will be able to reach. For sure, reaching goals makes you happy, but just enjoy the sport you love!
INVISIBLE HEDGEHOG (2):	**I think** that you can become an Olympian in RG if you move to...something like Marshall islands or other exotic countries and could get the spot as a Tripartile Commission of FIG-NOC-IOC like Cape Verde usually does)
USCOACH (3):	Simple answer "NO" You can have fun, but Olympic games are reserved for very few girls. Many girls who have been training for 15+ years don't even dare to dream of going to the Olympics. What you need to do is be realistic. People who end up going to the Olympics dedicate their entire life to the sport, have the right circumstances going for them, the right talent, the best coaching teams, the best support and they are also lucky...you can be an amazing gymnast but if it's not your time when the Olympic games come, you may miss your chance. To many factors to consider. Honestly **I think** your question is ridiculous, and borderline disrespectful.
TAHNEE (1):	I'm kind of glad someone's said this. It sounds harsh, but it's true. Take me for example - I've trained for nearly 17 years, I started at age 4. But I will never be going to the Olympics, even though I do work hard at training and have throughout my entire life in the sport. I had natural talent, and a very flexible back, but I was always realistic – the Olympics is something you need to train 30-40 hours a week for, for a great deal of your time in RG, as well as having natural talent, a good coach, good training conditions, and having the RG politics in your country in your favour is also

helpful. Like someone mentioned previously, it's also very difficult to pick up the body and apparatus technique at age 19, not to mention the fact that you're fully grown by then and your flexibility is generally limited. Although I have a flxible back when I was younger, and trained regularly, my flex has decreased dramatically over the last 5 years. It's hard enough to maintain when you're in the sport for your whole life, let alone if you just blow in at 19. So no, you won't become an Olympian, but if you love it that much, there's no reason why you shouldn't give it a go.

References

Aijmer, Karin

1997 *I think* – an English modal particle. In *Modality in Germanic Languages: Historical and Comparative Perspectives*, Toril Swan, and Olaf Jansen Westvik (eds.), 1—47. Berlin/New York: Mouton de Gruyter.

2009 Seem and evidentiality. *Functions of Language 16* (1): 63—88.

Aikhenvald, Alexandra

2004 *Evidentiality*. Oxford: Oxford University Press.

Androutsopoulos, Jannis

2006 Introduction:Sociolinguistics and computer - mediated communication. *Journal of Sociolinguistics 10* (4): 419—438.

Bakhtin, Mikhail M.

1981 *The Dialogic Imagination: Four Essays*. Caryl Emerson, and Michael Holquist (trans.), Michael Holquist (ed.). Austin: University of Texas Press.

Baron, Naomi S.

2010 Discourse structures in Instant Messaging: The case of utterance breaks. *Language@Internet 7* (4).

Biber, Douglas, and Edward Finegan

1989 Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text* 9 (1): 93-124.

Biber, Doughlas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan

1999 *Longman Grammar of Spoken and Written English*. Harlow: Longman.

"Bulletin board, n." OED Online. June 2013. Oxford University Press. 19 June 2013.

Bybee, Joan

1985 *Morphology*. Amsterdam: John Benjamins

Bybee, Joan, and Paul Hopper (eds.)

2001 *Frequency and the Emergence of Linguistic Structure*. Amsterdam/Philadelphia: John Benjamins.

Capelli, Gloria

2007 *"I Reckon I Know how Leonardo Da Vinci must have Felt…": Epistemicity, Evidentiality and English Verbs of Cognitive Attitude*. Pari, Italy: Pari Publishing.

Chafe, Wallace
1986 Evidentiality in English conversation and academic writing. In *Evidentiality: The Linguistic Coding of Epistemology,* Wallace Chafe, and Johanna Nichols (eds.), 261—272. Norwood, NJ: Ablex Publishing Corporation.

Claridge, Claudia
2007 Constructing a corpus from the web: message boards. In *Corpus Linguistics and the Web,* Marianne Hundt, Nadja Nesselhauf, and Carolin Biewer (eds.), 87—108. Amsterdam/New York: Rodopi.

Cornillie, Bert
2009 Evidentiality and epistemic modality: On the close relationship between two different categories. *Functions of Language 16* (1): 44-62.

Cruse, Alan
2004 *Meaning in Language*. Oxford: Oxford University Press.

Crystal, David
2001 *Language and the Internet.* 2d ed. New York: Cambridge University Press.

Davis, Boyd H., and Jeutonne P. Brewer
1997 *Electronic Discourse: Linguistic Individuals in Virtual Space.* Albany, NY: State University of New York Press.

Divjak, Dagmar
2006 Ways of intending; A corpus-based cognitive linguistic approach to near-synonyms in Russian. In *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis*, Stefan Th. Gries, and Anatol Stefanowitsch (eds.), 19—56. Berlin/New York: Mouton de Gruyter.
2010 *Structuring the Lexicon: A Clustered Model for Near-Synonymy.* Berlin/New York: Mouton de Gruyter.

Du Bois, John W.
2007 The stance triangle. In *Stancetaking in Discourse,* Robert Englebretson (ed.), 69—110. Amsterdam/Philadelphia: John Benjamins.

Englebretson, Robert. (ed.)
2007 *Stancetaking in Discourse.* Amsterdam/Philadelphia: John Benjamins.

Fillmore, Charles J.
1985 Frames and the semantics of understanding. *Quaderni de Semantics* 6: 222—254.

Ford, Cecilia E.
1997 Speaking conditionally: some contexts for *if*-clauses in conversation. In *On Conditionals Again*, Angeliki Athanasiadou, and René Dirven (eds.), 387—413. Amsterdam/Philadelphia: John Benjamins.

Ford, Cecilia E., and Sandra A. Thompson
1996 Interactional units in conversation. In *Interaction and Grammar*, Elinor Ochs, Emanuel A. Schegloff, and Sandra A. Thompson (eds.), 134—184. Cambridge/New York: Cambridge University Press.

Fortescue, Michael
2001 Thoughts about thought. *Cognitive Linguistics* 12 (1): 15—45.

Geeraerts, Dirk, Stefan Grondelaers, and Peter Bakema
1994 *Structure of Lexical Variation: Meaning, Naming, and Context.* Berlin/New York: Mouton de Gruyter.

Geeraerts, Dirk, Gitte Kristiansen, and Yves Peirsman (eds.)
2010 *Advances in Cognitive Sociolinguistics*. Berlin/New York: Mouton de Gruyter.

Gibbs, Raymond, and Jennifer O'Brien
1991 Psychological aspects of irony understanding. *Journal of Pragmatics 16* (6): 523—530.

Giora, Rachel
1995 On irony and negation. *Discourse Processes 19* (2): 239—264.

Glynn, Dylan
2009 Polysemy, syntax, and variation: A usage-based method for Cognitive Semantics. In *New Directions in Cognitive Linguistics*, Vyvyan Evans, and Stephanie Pourcel (eds.), 77—106. Amsterdam/Philadelphia: John Benjamins.
2010a Corpus-driven Cognitive Semantics: Introduction to the field. In *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*, Dylan Glynn, and Kerstin Fischer (eds.), 1—41. Berlin/New York: Mouton de Gruyter.
2010b Testing the hypothesis: Objectivity and verification in usage-based cognitive semantics. In *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches,* Dylan Glynn, and Kerstin Fischer (eds.), 239—269. Berlin/New York: Mouton de Gruyter.

in press Correspondence Analysis: An exploratory technique for identifying usage patterns. In *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*, Dylan Glynn, and Justyna Robinson (eds.). Amsterdam: John Benjamins.

Gries, Stefan Th.

2003 *Multifactorial Analysis in Corpus Linguistics: A Study of Particle Placement.* London: Continuum Press.

Hanley, James A., and Barbara J. McNeil

1982 The meaning and use of the area under a receiver operator characteristic (ROC) curve. *Radiology 143* (1): 19—36.

Herring, Susan C.

2002 Computer-mediated communication and the Internet. *Annual Review of Information Science and Technology* 36 (1): 109—168.

2003 Gender and power in online communication. In *The Handbook of Language and Gender*, Janet Holmes, and Miriam Meyerhoff (eds.), 202—228.

2004 Computer-mediated discourse analysis: An approach to researching online communities. In *Designing for Virtual Communities in the Service of Learning*, Sasha A. Barab, Rob Kling, and James H. Gray, (eds.), 338—376. Cambridge: Cambridge University Press.

Holmes, Janet

1990 Hedges and boosters in women's and men's speech. *Language & Communication* 10 (3): 185—205.

Hopper, Paul

1987 Emergent grammar. *BLS* 13: 139—155.

Hunston, Susan

2007 Using a corpus to investigate stance quantitatively and qualitatively. In *Stance-Taking in Discourse: Subjectivity, Evaluation, Interaction*, Robert Englebretson (ed.), 27—48. Amsterdam/Philadelphia: John Benjamins.

Hunston, Susan, and Geoffrey Thompson, Geoffrey (eds.)

2000 *Evaluation in Text*. New York. Oxford University Press.

Hård af Segerstad, Ylva

2002 Use and adaptation of written language to the conditions of computer-mediated communication. Ph.D. diss., Department of Linguistics, Göteborg University.

Jakobson, Roman
1960 Concluding statement: Linguistics and poetics. In *Style in Language*, Thomas A. Sebeok (ed.), 350—377. New York: John Wiley & the Sons and The Technology Press of Massachusetts Institute of Technology.

Kaltenböck, Gunther
2013 The development of comment clauses. In *The Verb Phrase in English*, Bas Aarts, Joanne Close, Geoffrey Leech, and Sean Wallis (eds.), 286—317. Cambridge: Cambridge University Press.

Kolko, Beth, and Elizabeth Reid
1998 Dissolution and fragmentation: Problems in online communities. In *Cybersociety 2.0: Revisiting computer mediated communication and community*, Steve Jones (ed.), 212—229.

Krawczak, Karolina and Dylan Glynn
2011 Context and cognition. A corpus-driven approach to parenthetical uses of mental predicates. In *Cognitive Processes in Language*, Janusz Badio, and Krzysztof Kosecki (eds.). Frankfurt/Main: Lang.

Kristiansen, Gitte, and René Dirven (eds.)
2008 *Cognitive Sociolinguistics: Language Variation, Cultural Models, Social System*. Berlin/New York: Mouton de Gruyter.

Kärkkäinen, Elise
2003 *Epistemic Stance in English Conversation: A Description of its Interactional Functions, with a Focus on* I think. Amsterdam/Philadelphia: Benjamins.
2006 Stance taking in conversation: From subjectivity to intersubjectivity. *Text & Talk 16* (6): 699—731.
2007 The role of *I guess* in conversational stancetaking. In *Stancetaking in discourse: Subjectivity, Evaluation, Interaction*, Robert Englebretson (ed.), 183—220. Amsterdam/Philadelphia: John Benjamins.

Lakoff, George
1987 *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

Lampert, Martin D., and Susan M. Ervin-Tripp
1993 Structured coding for the study of language and social interaction. In *Talking Data: Transcription and Coding in Discourse Research*, Jane A. Edwards, and Martin D. Lampert. Hillsdale, NJ: Lawrence Erlbaum Associates.

Langacker, Ronald, W.
1987 *Foundations of Cognitive Grammar.* Vol. 1. *Theoretical Prerequisites*. Stanford: Stanford University Press.
1991 *Foundations of Cognitive Grammar*. Vol 2. *Descriptive Application*. Stanford: Stanford University Press.
Lattin, James M, Douglas J. Carrol, and Paul E. Green
2003 *Analyzing Multivariate Data*. Pacific Grove, CA: Thomson Brooks.
Lee, Jennifer
2002 I think, therefore IM. In *New York Times*, September 19, Circuits Section, 1.
Lewin, Berverly A., and Yonatan Donner
2002 Communication in Internet message boards: A quantitative analysis of usage in Computer-Mediated Conversation (CMC). *English Today 18* (3): 71—98.
Linell, Per
1995 Troubles with mutualities: Towards a dialogical theory of misunderstanding and miscommunication. In *Mutualities in Dialogue*, Ivana Marková, Carl Graumann, and Karl Foppa (eds.), 176—213. Cambridge: Cambridge University Press.
1998 *Approaching Dialogue: Talk, Interaction and Contexts in Dialogical Perspectives*. Amsterdam/Philadelphia: John Benjamins.
2009 *Rethinking Language, Mind, and World Dialogically: Interactional and Contextual Theories of Human Sense-Making.* Charlotte/NC: Information Age Publishing.
Lyons, John
1977 *Semantics*. Cambridge: Cambridge University Press.
Marcoccia, Michel
2004 On-line polylogues: conversation structure and participation framework in internet newsgroups. *Journal of Pragmatics* 36 (1): 115—145.
Marková, Ivana
2003 *Dialogicality and Social Representations: The Dynamics of Mind*. Cambridge: Cambridge University Press.
Marková, Ivana, Per Linell, Michele Grossen, and Anne Salazar Orvig
2007 *Dialogue in Focus Groups: Exploring Socially Shared Knowledge.* London: Equinox.
Martin, James R., Peter R.R. White
2005 *The Language of Evaluation: Appraisal in English*. New York: Palgrave Macmillan.

Nuyts, Jan
2001 *Epistemic Modality, Language, and Conceptualization: A Cognitive-Pragmatic Perspective.* Amsterdam/Philadelphia: John Benjamins.

Palmer, Frank R.
1986 *Mood and Modality*. Cambridge: Cambridge University Press.
2001 Reprint. *Mood and Modality*. 2d ed. Cambridge: Cambridge University Press. Original edition, Cambridge: Cambridge University Press, 1986.

Postmes, Tom, Russell Spears, Khaled Sakhel, and Saphne de Groot
2001 Social influence in computer-mediated communication: The effects of anonymity on group behavior. *Personality and Social Psychology Bulletin Board 27* (10): 1242—1254.

Reid, Elizabeth
1999 Hierarchy and power: Social control in cyberspace. In *Communities in Cyberspace*, Marc A. Smith, and Peter Kollock (eds.), 107—133.

Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson
1974 A simplest systematics for the organization of turn-taking for conversations. *Language 50* (4): 696—735.

Scheibman, Joanne
2002 *Point of View and Grammar: Structural Patterns of Subjectivity in American English Conversation.* Amsterdam/Philadelphia: John Benjamins.

Schmid, Hans-Jörg
1993 *Cottage and Co., Idea, Start vs. Begin: Die Kategorisierung als Grundprinzip einer differenzierten Bedeutungsbeschreibung*. Tübingen: Niemeyer.

Spears, Russell, and Martin Lea
1994 Panacea or panopticon? The hidden power in computer-mediated communication. *Communication Research 21* (4): 427—459.

Speelman, Dirk
in press Logistic Regression: A confirmatory technique for comparisons in corpus linguistics. In *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*, Dylan Glynn, and Justyna Robinson (eds.). Amsterdam: John Benjamins.

Speelman, Dirk, and Dirk Geeraerts
2010 Causes for causatives: The case of Dutch ‘doen’ and ‘laten’. In *Causal Categories in Discourse and Cognition*, Ted Sanders, and Eve Sweetser (eds.), 173—204. Cambridge: Cambridge University Press.

Sperber, Dan, and Deirdre Wilson
1981 Irony and the use-mention distinction. In *Radical Pragmatics*, Peter Cole (ed.), 295—318. New York: Academic Press.
Stefanowitsch, Anatol, and Stefan Th. Gries
2003 Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics 8* (2): 209—243.
Sweetser, Eve
1990 *From Etymology to Pragmatics*. Cambridge: Cambridge University Press.
Thompson, Sandra A., and Anthony J. Mulac
1991 A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In *Grammaticalization II*, Elizabeth Traugott, and Bernd Heine (eds.), 313—339. Amsterdam/Philadelphia: John Benjamins.
Tomasello, Michael
2003 *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
Van der Auwera, Johan, and Vladimir Plungian
1998 Modality's semantic map. *Linguistic Typology* 2 (1): 125—139.
Vendler, Zeno
1957 Verbs and times. *The Philosophical Review* 66 (2): 143—160.
Verhagen, Arie
2005 *Constructions of Intersubjectivity: Discourse, Syntax, and Cognition*. Oxford: Oxford University Press.
Vološinov, Valentin
1986 Reprint. *Marxism and the Philosophy of Language.* Cambridge, MA: Harvard University Press. Original edition, Cambridge, MA: Harvard University Press, 1973.
Werry, Christopher C.
1996 Linguistic and interactional features of Internet Relay Chat. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, Susan Herring (ed.), 47—63. Amsterdam/Philadelphia: John Benjamins.

Wulff, Stefanie, Anatol Stefanowitsch, and Stefan Th. Gries

2007 Brutal Brits and persuasive Americans: Variety-specific meaning construction in the into-causative. In *Aspects of Meaning Construction*, Günter Radden, Klaus-Michael Köpcke, Thomas Berg, and Peter Siemund (eds.), 265—281. Amsterdam/Philadelphia: John Benjamins.