

Vilnius University
Faculty of Philology
Department of English Philology

Roberta Ambraškaitė

A Contrastive Analysis of Vague Language in Spoken English of Lithuanian Learners of English and Native Speakers

Thesis submitted in partial fulfilment of requirements
for the degree of MA in English Studies

Supervisor: Dr Rita Juknevičienė

2016

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr R. Juknevičienė for the guidance throughout this study as well as valuable suggestions and insight. I also wish to thank my supervisor for continuous encouragement and support. Moreover, thanks to Dr Birutė Ryvitytė for a very helpful course on Research methodology and my course mates whose comments provided during the seminars were of great value. All the errors that remain, however, are my own.

CONTENTS

| | |
|---|----|
| Abstract..... | 5 |
| 1. Introduction | 6 |
| 1.1. General introduction to VL..... | 6 |
| 1.2. Overview of the relevant concepts..... | 8 |
| 1.3. Classification of vagueness markers..... | 11 |
| 1.4. Review of previous research..... | 14 |
| 2. Data and Methods..... | 18 |
| 3. Results and Discussion | 21 |
| 3.1. Quantitative analysis..... | 21 |
| 3.1.1. Underused vagueness markers in LINDSEI-LITH | 22 |
| 3.1.2. Overused vagueness markers in LINDSEI-LITH | 22 |
| 3.1.3. Frequencies of vagueness markers with insignificant statistical differences..... | 23 |
| 3.2. Qualitative analysis..... | 24 |
| 3.2.1. Vague coordination tags..... | 24 |
| 3.2.1.1. Structural analysis of underused tags in LINDSEI-LITH | 27 |
| 3.2.1.2. Functional analysis of underused tags in LINDSEI-LITH..... | 30 |
| 3.2.1.3. Structural analysis of overused tags in LINDSEI-LITH | 33 |
| 3.2.1.4. Functional analysis of overused tags in LINDSEI-LITH..... | 34 |
| 3.2.2. Compromiser <i>quite</i> | 35 |
| 3.2.2.1. Structural analysis of <i>quite</i> | 37 |
| 3.2.2.2. Functional analysis of <i>quite</i> | 42 |
| 3.2.3. Compromiser <i>rather</i> | 45 |
| 3.2.3.1. Structural analysis of <i>rather</i> | 47 |
| 3.2.3.2. Functional analysis of <i>rather</i> | 49 |
| 3.2.4. Compromiser <i>kind of</i> | 49 |
| 3.2.4.1. Structural analysis of <i>kind of</i> | 53 |

| | |
|--|----|
| 3.2.4.2. Functional analysis of <i>kind of</i> | 58 |
| 4. Conclusions | 61 |
| 5. Summary in Lithuanian | 64 |
| References | 66 |
| Appendices | 69 |

Abstract

Vagueness and vague language (VL) is an inseparable part of natural language. Furthermore, spontaneous spoken language, conversation in particular, relies on non-elaboration, dysfluencies, simplicity etc. all of which contribute to naturalness of speech. Among such features, linguists have also identified vague language which, as research shows, is often misused or underused by non-native learners of English, which, as a consequence, leads to artificially sounding speech. This study aims to investigate VL in native and non-native spoken English in order to answer the question why EFL learners sound unnatural when compared to native speakers. Two corpora, LINDSEI-LITH and LOCNEC, were chosen for the comparative analysis of learners' and native speakers' use of different vagueness markers (VMs). Firstly, the quantitative analysis was done, which revealed that there are more underused than overused VMs in EFL learner English. Secondly, the qualitative analysis of the most often used vague coordination tags and compromisers in LINDSEI-LITH was done. The analysis was twofold: structural and functional, and thus revealed not only syntactic patterns and lexical combinability but also similarities and differences in the functions of VMs.

1. Introduction

1.1. General introduction to VL

Perhaps one of the first disciplines in which the interest in vagueness arose is philosophy. The Greeks were concerned about so-called *sorites paradox*, which points to the problem of having clear boundaries. One of the best-known examples of a sorites paradox is the term ‘heap’, which is vague and the boundary between *a heap* and *a non-heap* cannot be identified as well as questions like ‘Does one grain of wheat make a heap? Do two grains of wheat make a heap? Do three grains of wheat make a heap?... Do ten thousand grains of wheat make a heap?’ cannot be answered (Williamson 1994, Overstreet 2011). Apart from philosophers, psychologists have also been interested in vagueness. They often claim that the problem of vagueness should be analysed by psychologists not linguists, because ideas that are expressed by language are vague, not the language itself. However, the question ‘what is vague’ is too difficult to answer and in contemporary world, we see both linguists and literary critics do research in vagueness. Drave (2002) notes that while philosophers deal with logical-language conception of what is vague, linguists work with a natural one. The term *vagueness* itself is rather vague and ambiguous (Ullmann 1962: 118) and the limits of vagueness are themselves vague (Williamson 1994), which makes the analysis of vagueness very complicated in any of the sources.

Peirce (1902) (in Channell 1994) is often referred to as the originator of the notion of vagueness in language. According to Channell, Peirce was perhaps the first one who attempted to formulate this notion in a very detailed way:

A proposition is vague where there are possible states of things concerning which it is intrinsically uncertain whether, had they been contemplated by the speaker, he would have regarded them as excluded or allowed by the proposition. By intrinsically uncertain we mean not uncertain in consequence of any ignorance of the interpreter, but because the speaker's habits of language were indeterminate (...) (Peirce 1902: 748, quoted in Channell 1994: 7)

By referring to propositional vagueness Peirce focuses on a language user and the meaning he or she wants to convey. Defining a vague expression, Channell refers to this definition: ‘[a]n expression or word is vague if: a) it can be contrasted with another word or expression which appears to render the same proposition; b) it is purposely and unabashedly vague; c) its meaning arises from the ‘intrinsic uncertainty’ referred to by Peirce (ibid. p. 20). Drave defines the same notion in a simpler way, claiming that one of the possible ways to define vague language (hereafter VL) is ‘as that which modifies a linguistic item, phrase or utterance to make its meaning less precise’ (2002: 25). For his study, in order to describe a closed set of

items that are considered as instances of vagueness, Drave also takes Powell's (1985) notion of being inherently or purposively vague and adds that it is impossible to paraphrase the meaning of such vague expressions in a non-vague way.

The notion of VL can also be found in grammar books. For example, Coffin et al. note that VL '[s]oftens expressions so that they do not appear too direct or unduly authoritative and assertive' (2004: 31). They also add that a more precise term would be *purposefully vague language*, because vague expressions sometimes are considered just a sign of carelessness. Similarly, Zhang (2015) also suggests modifying the terminology. He notes that the term *vague* has a negative connotation and replaces it by *elastic*, aiming to provide a positive representation of the language of such nature. However, for the purpose of this study, I chose the terminology and functions of VL, which is still more conventional.

One of the beliefs about 'good' language is that it has to be clear and precise; however, Channell (1994) argues that such beliefs and the instructions how to write as the ones provided by Partridge, who was promoting good English usage, noting that '[a person should write] so clearly, so precisely, so unambiguously, that his words can bear only one meaning to all averagely intelligent readers (...)' (1947: 437, quoted in Channell 1994: 1), can mislead especially those who use English as a second language. Channell notes that VL cannot be treated as all 'bad' or all 'good', because through the use of a degree of vagueness writers demonstrate their competence, if the usage is 'right for the purpose of their writing' (ibid. p. 3). Therefore, before making some judgements about VL, the usage of it should be observed closely. Interestingly, Wittgenstein (1953, quoted in Channell 1994: 6) was already claiming that often vagueness is what is needed. Moreover, Biber et al. (1999: 113) in their corpus-based grammar also assert that approximating expressions are used 'where complete explicitness may not be necessary, and some degree of vagueness may actually be desirable'. They also add that in such contexts greater precision would be superfluous, would need more processing, and thus the dynamic of conversation would be consequently delayed (ibid. p. 1045).

Although there can be attempts to avoid vagueness in academic discourse, where language is rather formal, clear and unambiguous, it is very difficult if not impossible to avoid it in more informal discourses such as speech, because it is a feature of naturally spoken authentic language. Thus, vagueness is one of the factors creating naturalness of expression and posing a major problem to a non-native speaker (NNS), because the very logic of language acquisition rests upon accuracy and clarity of expression, while vagueness is anything but not clarity and precision of expression. This, naturally, makes VL unattractive to NNS users,

prompting learners to avoid such expressions, which leads to producing artificially sounding speech. Therefore, it is necessary to investigate VL in spoken discourse; however so far most of the studies have focused on the usage of VL in written discourse, thus in this paper I will focus on spoken English. Hence, before concentrating merely on VL, other relevant concepts will be briefly overviewed in section 1.2., classifications of vagueness markers (VMs) in 1.3., and previous research in 1.4.

1.2. Overview of the relevant concepts

Firstly, three quite broad terms that can be seen as umbrella terms, *metadiscourse*, *stance* and *epistemic modality*, all of them including hedges as a narrower term, will be introduced. Then, the relation between *hedges* and VL will be addressed. Finally, the characteristics of *spoken discourse* will be considered.

To begin with, according to Hyland, ‘[t]he concept of metadiscourse is based on a view of writing (and speaking) as a social and communicative engagement, offering a means of understanding the ways we project ourselves into our texts to manage our communicative intentions’ (2005: 14). Thus, Hyland using this concept refers to the aspects of the text that ‘explicitly refer to the organisation of the discourse or the writer’s stance towards either its content or the reader’ (1998b: 438). The former might be referred to as interactive resources and the latter as interactional resources (Hyland and Tse 2004: 168). Both types of resources can be divided into narrower categories, but for this study only the interactional resources are relevant, and thus Table 1 presents categories, functions and examples of interactional resources.

Table 1. Metadiscourse: interactional resources (Hyland and Tse 2004: 169)

| Category | Function | Examples |
|---------------------------|--|---------------------------------------|
| Hedges | withhold commitment and open dialogue | might/ perhaps/ possible/ about |
| Boosters | emphasise certainty or close dialogue | in fact/ definitely/ it is clear that |
| Attitude markers | express writer's attitude to proposition | unfortunately/ I agree/ surprisingly |
| Self-mentions | explicit reference to author(s) | I/ we/ my/ me/ our |
| Engagement markers | explicitly build relationship with reader consider | note/ you can see that |

As Table 1 shows interactional resources include hedges and boosters, which are most relevant categories to my study, as well as attitude markers, self-mentions and engagement markers. Interestingly, these five categories can also be arranged in a different way in order to explain the reader-writer interaction (see Fig.1)

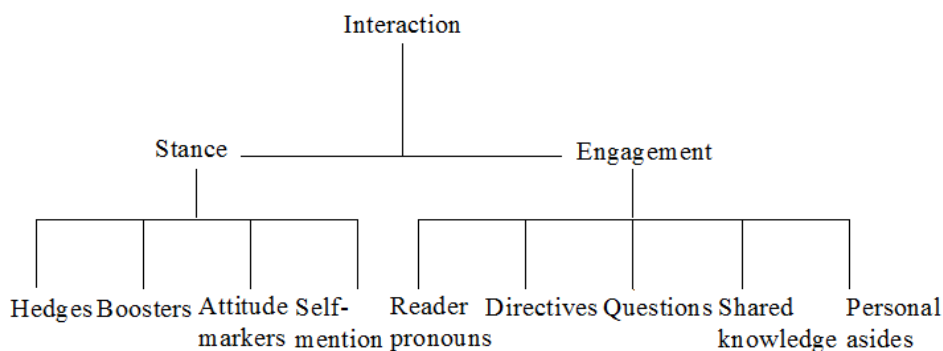


Fig.1. Key resources of academic interaction (Hyland 2005a: 177)

Hyland (2005a) notes that both *stance* and *engagement* contribute to interpersonal dimension of discourse, thus there are overlaps between them. *Stance* refers to writers’ presentations of themselves as well as conveyance of their judgements, opinions, and commitments, and thus can be seen as an attitudinal dimension. *Engagement*, on the other hand, can be seen as an alignment dimension, in which writers connect to the readers, ‘pulling them along with their argument, focusing their attention, acknowledging their uncertainties, including them as discourse participants, and guiding them to interpretations’ (Hyland 2005a: 176).

Another important concept is *modality*. Dealing with hedging, *epistemic modality* is most relevant, while *non-epistemic modality* is important only for the distinction of these two types, as modality markers can have several different functions (Šinkūnienė 2011: 29). According to Hyland (1998a: 44), ‘[e]pistemic modality expresses the speaker’s opinion or belief concerning the truth of what is said’, or in other words, it displays confidence, e.g. through the use of boosters such as *certainly*, or lack of confidence, e.g. using hedges such as *perhaps*, regarding the truth of propositional information. Discussing epistemic modality we can refer to possibility, probability, or inferred certainty. Thus, epistemic modality is mostly related with author’s judgement about the truth of the proposition (Usonienė 2004: 40).

Different scholars treat VL differently and due to the different understanding of the same concepts, some researchers tend to point out the close link between VL and hedges, while the others choose not to discuss this issue. The first one to use the term hedges was Lakoff: ‘(...) words whose meaning implicitly involves fuzziness – words whose job is to make things fuzzier or less fuzzy. I will refer to such words as ‘hedges’’(1973: 471). According to Overstreet (2011), Lakoff’s analysis provided the impetus for an expanding area of investigation identifying VL as one of the main devices used for hedging. Thus, there is a common view that VL is a sub-category of hedging that was also noted by Šinkūnienė (2011) and Ruzaitė (2004). Drave takes this view for his own study and defends it by claiming that ‘VL is more specifically concerned with propositional precision or exactness’ (2002: 26).

Moreover, Drave claims that both, vagueness and hedging, are found within the domain of epistemic modality, because they convey ‘the relationship between a speaker or writer and the matter of an utterance or statement’ (ibid.). While Channell (1994) does not refer to the relation between these two modality markers at all, but as Šinkūnienė notes, some intentions of using VL that are given by Channell, in fact, overlap with the ones of hedging, for instance, being polite, trying to protect yourself from criticism or when the specific information is unknown to the speaker (2011: 41).

Finally, as spoken discourse is at the core of this study, the characteristics of it will be considered. Biber et al. (1999: 1041) emphasise that the primary function of conversation is ‘to establish and maintain social cohesion through the sharing of experience’, while it may have the secondary function of entertainment, e.g. through narratives, or exchange of information or even control the behaviour of other people. In a summary of characteristics of speech and writing discussed by a number of linguists and provided by Biber (1988: 47), it is noted that speech depends more on shared situation and background knowledge and is more personally involved than writing. Moreover, the concentration of new information is lower and it is less deliberately organized and planned. Needless to say, it is also spontaneous, thus there is not much time for planning and choosing the right words. Repetition is another feature of a conversation, which is used either to emphasise something or to make something clear (Biber et al. 1999:53).

Brown and Yule (1983: 13) added some information on the syntax of spoken language: it contains many incomplete sentences as well as little subordination. Biber (1988: 47) also noted that speech is less structurally complex and elaborate as well as explicit than writing. For example, usually there are maximum two premodifying adjectives and the forms that are found are active and declarative. Moreover, while metalingual markers to mark relationships between clauses are used in written language, in spoken language chunks are linked by *and*, *but*, *then* and, *if*. To add more, one of the characteristics that perhaps could not be applied to written language in any case is that ‘the speaker may rely on [e.g.] gaze direction to supply a referent’ (Brown and Yule 1983: 15-17). Consider the following examples (ibid.):

(1) (looking at the rain) *frightful isn't it*

In addition, expressions may be replaced or refined by the speaker, which could perhaps sometimes help to avoid vagueness, e.g.:

(2) *this man + this chap she was going out with*

As regards vocabulary, it could be said that it is rather generalised, e.g. *a lot of, things like that*, and a huge number of prefabricated ‘fillers’ such as *you know, if you see what I mean* are used by the speakers. It should also be mentioned that speech is believed to contain most vague expressions. De Cock mentions that Crystal and Davy (1975), Aijmer (2002), Channell (1994) and Drave (2002) concluded that ‘lack of precision is one of the most important features of informal interaction’ (2004: 236). Four reasons why vagueness occurs provided by Crystal and Davy (in Channell 1994: 8) are: 1) memory loss, 2) the language has no suitable exact word, or the speaker does not know it, 3) the subject of the conversation is not such that it requires precision, and an approximation or characterization will do, 4) the choice of a vague item is deliberate to maintain the atmosphere. Channell (ibid. p. 4) also states that ‘normally we do not notice [VL] unless it appears inappropriate’, such as deliberately withholding information. Moreover, misused vague expressions would also be noticed and would indicate the speaker being non-native.

Few classifications of vagueness markers will be presented in section 1.3.

1.3. Classifications of vagueness markers

Markers of vagueness (VMs) can be classified in few ways. In this section three classifications that are provided by Biber et al. (1999), Channell (1994), and Quirk et al. (1985), and are relevant to my study will be presented.

To begin with, Channell (1994) distinguishes and analyses three main classes of vague language: vague additives, vagueness by choice of vague words and vagueness by implicature. See Fig. 2 for explicit presentation of these classes:

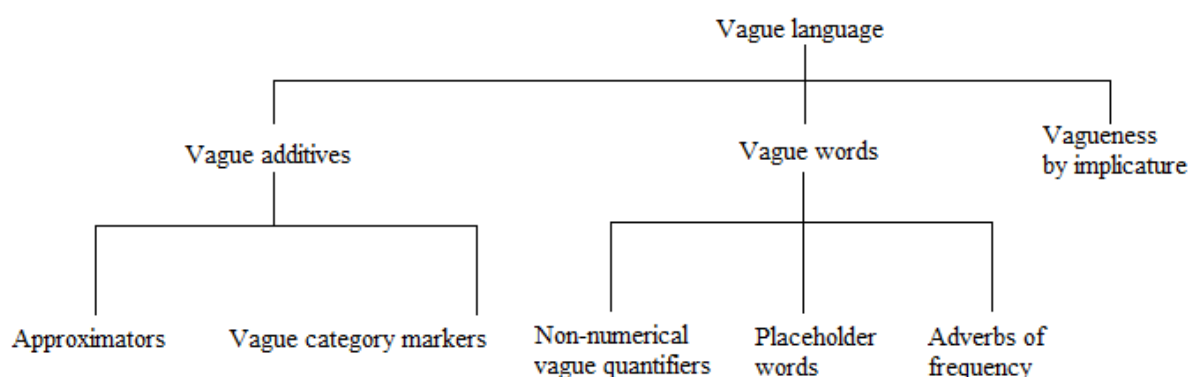


Fig.2. The classification of VL by Channell (1994)

According to Channell, vague additives are those words or phrases that are added to a precise statement and make it vague, for example the number approximations *about, around, round, approximately*, which are required to be followed by something exact. The exact numbers are

also vague if they are used in the constructions like *n or m, n or so*¹. The latter is a clear example of vague category markers, also called tags, which include phrases like *or something (like that), or anything (like that), and so on*. To achieve vagueness by choice of vague words, speakers choose words that are always vague, for instance, non-numerical vague quantifiers like *heaps of, a touch of, lots, some, several*, or placeholder words that do not belong to the standard written English and replace names or item names as *thingy, whatsisname* or *whatnot*. The third type of vagueness, vagueness by implicature, refers to cases in which a precise sentence can be understood as having a vague meaning. For example, the sentence *Sam is six feet tall* can be interpreted either as having precise meaning, claiming that Sam is exactly six feet tall, or as vague, concluding that he is approximately six feet tall (Channell 1994: 18-19). Jucker et al. (2003) took Channell's classification and for their own study added downtoners, e.g. *sort of, kind of*, and vague adverbs of likelihood such as *probably* and *maybe*. The classification of VL in conversation was also provided by Biber et al. (1999), in which they refer to two groups as different types of hedges. Consider Fig. 3:

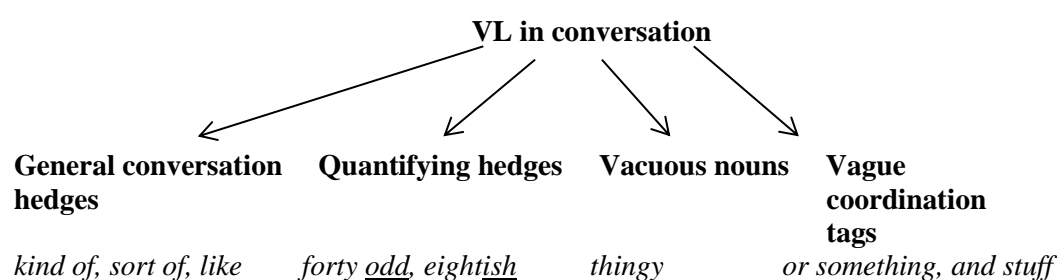


Fig.3. The classification of VL in conversation (Biber et al. 1999: 1045)

As the figure shows, VMs such as *kind of, sort of, like*, according to Biber et al. (ibid.), belong to the group of general conversation hedges, *odd of forty odd* and *eightish* are quantifying hedges. Moreover, words like *thingy* that were referred to by term placeholder words by Channell (1994), in this classification are referred as vacuous nouns. Finally, vague coordination tags, e.g. *or something*, are also present in this classification.

Although downtoners are not included in the main classification, in their corpus-based grammar Biber et al. noted that some downtoners, or degree adverbs scaling down the effect of the modified item, ‘convey some sense that the use of the modified item is not precisely accurate’ (1999: 555). Fig. 4 illustrates classification of intensifiers that include downtoners:

¹ *N* and *m* stand for numbers.

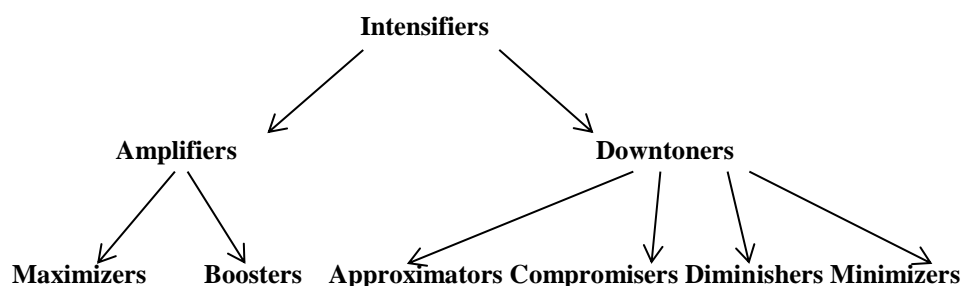


Fig.4. The classification of intensifiers (Quirk et al. 1985: 589-590)

Intensifiers indicate a point, which can be relatively low or high, on a scale. Thus, they can express not only an increase in intensification, but also a decrease. Amplifiers scale upwards from an assumed norm as in ‘a very funny film’ compared to ‘a funny film’, while downtoners have the opposite effect: they scale downwards from an assumed norm as in ‘It was almost dark’ compared to ‘It was dark’ (Quirk et al. 1958: 445-446). Amplifiers can be further divided into maximizers, which denote the upper extreme scale, e.g. *completely*, and boosters, which deal with a high point on the scale just as *very much* (ibid. 590-591). Downtoners are more relevant to my study, and thus Table 2 presents the functions of different types of downtoners and provides examples.

Table 2. Functions of downtoners (Quirk et al. 1958: 597-598)

| Type of downtoners | Function | Example |
|----------------------|--|------------------------------------|
| Approximators | serve to express an approximation to the force of the verb, while indicating that the verb concerned expresses more than is relevant | <i>almost, nearly, practically</i> |
| Compromisers | have only a slight lowering effect and tend to call in question the appropriateness of the verb concerned | <i>kind of, quite, rather</i> |
| Diminishers | scale downwards and roughly mean ‘to a small extent’ | <i>quite, partially, slightly</i> |
| Minimizers | they are negative maximizers, ‘(not) to any extent’ | <i>barely, hardly, scarcely</i> |

Quirk et al. also emphasise that compromisers, having an ability only to lower the force, ‘cannot deny the truth value of what is denoted by the verb’ (ibid. p. 600) as in *I kind of like him*. Thus, this example cannot be interpreted as meaning **in fact I don’t like him*, and the closest interpretation is actually: ‘I might go as far as to say I like him’ (ibid.). To add more, *kind of* and *sort of* are sometimes used as approximators with non-gradable verbs as in *He kind of grunted* or *He sort of smiled at us*. Quirk et al. suggest that in here you could almost say ‘he grunted’ and ‘he smiled at us’ or ‘he sort of smiled at us, but in fact it was more like a

sneer.’ Although in quite many cases we could paraphrase sentences using *almost*, for other speakers *kind of* is closer to *more or less* than *almost*. However, it is almost impossible to identify whether the phrase means *almost* or *more or less* to the speaker and in order to avoid being highly subjective, these different interpretations were not taken into consideration.

As it can be noted from the table, some VMs, e.g. *quite*, in different contexts can have different functions and be treated as a compromiser or diminisher; however, functions of VMs will be discussed later in the paper.

1.4. Review of previous research

Firstly, the brief overview of previous research related to VL in written discourse will be presented, which will be followed by the overview of the research in spoken discourse. Finally, the studies on learners’ language in relation to VL will be presented.

First of all, hedges in academic discourse have been studied by a number of authors, because they are significant for effective academic writing. For example, Hyland (1998) studied hedging and boosting in research articles from a number of different disciplines. Hyland (1994) also examined hedging in textbooks, focusing on the coverage of lexical items as markers of uncertainty and tentativeness. Moreover, the use of hedges by Lithuanian writers has also been investigated, e.g. Šinkūnienė (2011) compared hedging in Lithuanian and English research articles in the field of humanities and biomedical sciences. It was found that the Lithianian authors employ hedges less frequently than the English authors.

Channell (1994) mentions some early studies of VL, e.g. a corpus-based study concerning the approximation devices in printed texts done by Kennedy (1987), which showed that approximation is frequent, and the study of *or something* in spoken discourse conducted by Ball and Ariel (1978). The other study of another coordination tag was by Ward and Birner (1993), who discussed the semantics and pragmatics of *and everything* in utterances, claiming that the meaning of the phrase is non-compositional, which means that it is not the predictable sum of the meanings of its components. Another very interesting study was done by Prince, Frader, and Bosk (1982) (also mentioned in Channell 1994) who analysed hedges in doctors’ working in an intensive care unit speech and found that a hedge occurs every fifteen seconds.

There have also been some cases when researchers compared the use of VL in two types of discourse. For example, Ruzaitė (2004) explored the use of approximators in the British National Corpus. She contrasted the use of these VMs in academic discourse with other types of discourse: leisure, business and political. Ruzaitė argues that imprecision in academic discourse is unavoidable and the omission of approximators could even cause some problems

for communication. It was also found that approximators appear much less frequently in spoken than in written academic discourse and that the frequency of them is higher only in spoken leisure discourse.

An interesting field of language where more research should be done on VL is the language used by the learners of English (EFL learners). It is often the case that learners are not able to use VMs appropriately. As it was mentioned before, vagueness plays an important role in NS language, thus in order to help learners sound more natural this particular field should get more attention. There have been already some useful studies carried on VL in learners' speech. For example, De Cock (2004) analysed the recurrent sequences of two or more single words that native speakers of English and advanced French EFL learners tend to use in their language, focusing on the spoken discourse, from a quantitative and qualitative point of view in order to test the validity of Kjellmer's assumption 'that learners' building material is individual bricks rather than prefabricated sections' (1991: 124 in De Cock). In the qualitative part of the paper, exploring the functions of the preferred sequences under the category of interactional/interpersonal sequences she discusses markers of vagueness (vagueness tags, vague category identifiers or general extenders, and a set of sequences that contain the discourse items *sort of* and *kind of*). Her study proves that French learners of English, in fact, underuse and also misuse markers of vagueness. It is concluded that the learners' speech lacks a typical feature of NS informal interactions: '... toning down and weaving the right amount of imprecision and vagueness' (ibid. p.243).

Similarly, Drave (2002) found that native speakers (NSs) of English use more VL than NSs of Cantonese while speaking in English. He also presented the functional differences of the category marker and placeholder *stuff* between NSs of English and NSs of Cantonese, proving that the latter do not employ all functions for which this particular word can be used. The research conducted by Orfanò (2013) on the use of VL by learners of English at a Brazilian university and NSs also showed that NSs use a broader range of vague category markers with different functions than the non-native learners. The study also revealed different preferences of these two groups of speakers for the use of particular vague category markers. Moreover, it was also concluded that learners are not aware of all the functions and uses of vague category markers in spoken language.

Aijmer (2004) analysed learners' use of pragmatic markers, focusing on Swedish learners' spoken English. Among other pragmatics markers she also analysed such as *kind of*, *or something like that*, *stuff like that*, *or so*. She concluded that these markers were used for different reasons by learners and NSs. While NSs use markers of vagueness and uncertainty

for face-saving or being polite, learners use them to express uncertainty or hesitation. According to Aijmer, learners also use markers when they have some communication problems. Furthermore, in order to fill a space in conversation and to avoid silent pauses, learners cluster the markers. An interesting fact is that Swedish learners sound more uncertain than native speakers, because they frequently use the marker *I don't know* (ibid. p. 188).

According to Grigaliūnienė and Juknevičienė (2011), writing of Lithuanian learners of English has been researched in different ways, yet speech remains largely *terra incognita*. The earliest research focused on lexical, or collocational, errors and the learners' language became the object of the studies only after the first Lithuanian corpus of written English learner (LICLE) (Grigaliūnienė et al. 2008) and a pilot corpus of the Lithuanian component of LINDSEI (LINDSEI-LITH) were compiled. Thus, learner corpus research is still a very young field in Lithuania and more research should be done in the field. However, there have been few important studies conducted on Lithuanian EFL learner's language. As regards written discourse, Šeškauskienė (2008) researched the use of hedging devices in BA paper introductions written by Lithuanian EFL learners, majoring in English. Jankauskaitė-Jokūbaitienė (2013) investigated boosters in argumentative essays by NSs and NNSs, including Lithuanian learners, in her MA thesis.

Moreover, Grigaliūnienė and Juknevičienė (2011) researched recurrent word phrases in the Lithuanian learner speech. It was found that the advanced Lithuanian learners of English use a number of formulaic sequences, but such sequences usually are not idiomatic. Thus, the formulaicity of most of the formulaic sequences in the Lithuanian learners' language is largely determined by pragmatic functions. Moreover, according to Grigaliūnienė and Juknevičienė, '(...) uncertainty and hesitation is a dominant feature of learner speech' (ibid. p. 16) and the results of their research show that formulaic sequences *I think, you know, I don't know* indicate uncertainty and hesitation of Lithuanian EFL learners' speech. In addition, Grigaliūnienė and Juknevičienė (2013) explored the recurrent formulaic sequences in the spoken and written English of Lithuanian EFL learners, showing that their speech is more formulaic than writing. Interestingly, a number of identical formulaic sequences occur in both language modes, which suggests that the learners have not fully acquired register awareness.

Although in the previous research there are some references to the ways Lithuanian EFL learners express their uncertainty and hesitation, there has not been research done focusing on the usage of VL. Moreover, in order to find out whether VMs are underused, overused or misused by Lithuanian learners, the findings should be compared to the NS usage of VL. Thus, the aim of this paper is to analyse VL used by NSs and Lithuanian learners of English

in spoken English from a quantitative point of view in order to find out the differences of the frequency of VMs in their speech. The second aim of this study is a qualitative analysis of VL in order to explore the tendencies found in the language of Lithuanian speakers and contrast findings with the usage by NSs as well as insights of previous research done on EFL learners having other mother tongue backgrounds. However, it can already be hypothesised that Lithuanian learners, similarly to other learners of English, are underusing VL and are not using full range of VL functions, and thus the functions differ comparing to NS language. It can also be hypothesised that learners avoid those VMs that do not have literal translational equivalents in their mother tongue. The study is expected to shed more light on the issue of naturalness of expression and, possibly, answer the question why EFL learners sound unnatural when compared to NSs.

2. Data and methods

There are two primary sources for my study: a pilot Lithuanian component of the Louvain International Database of Spoken English Interlanguage (LINDSEI-LITH, Grigaliūnienė and Juknevičienė 2011), and the Louvain Corpus of Native English Conversation (LOCNEC, De Cock 2004). The first corpus consists of 79 informal interviews, 90 176 words in total, with advanced learners of English, who were third-year students of English philology at the time of recording, all of them NSs of Lithuanian. Two NSs of English conducted the interviews, each lasting approximately 15 minutes. LOCNEC is a comparable corpus of 50 interviews with NSs, 146 121 words. This corpus is specifically compiled to be used in contrastive studies of learner language. Thus, both corpora follow the same pattern: firstly, students discuss their university life, hobbies, travel or plans for the future; then, they speak about an experience that taught them an important lesson, a country which impressed them or a film/play they particularly liked or disliked; finally, students have to tell a short story based on a picture.

Each interview has a unique code, which is composed of letters and numbers. *LT* shows that the extract is produced by a Lithuanian-speaking learner, while *ENG* signifies that it is by a native interlocuter. A number following the letters indicates the number of the interview in the corpus. Moreover, all transcribed interviews are divided into <A>, indicating the questions and comments given by the interviewer, and , indicating the answers by the interviewee, turns. For this study only turns were taken.

To avoid any confusion the term *vagueness marker* (VM) will be used to refer to any vague expression or word in this paper. Thus, a list of 52 VMs based on the previous research (Channell, 1994; Jucker et al. 2003; De Cock 2004) as well as the information about VL from grammar books (Quirk et al. 1985; Biber et al. 1999) was compiled. Using *Antconc* (Anthony 2012), a concordance program, all VMs in the context were extracted from the corpora and sorted by frequency of occurrence. The alphabetical list of these VMs can be seen below (see Appendix 1 for raw frequencies):

(A) few, a bit of, a couple of, a little, a little bit, a lot of, about, almost, and all that, and everything, and so on, and stuff, and things, anywhere, approximately, around, et cetera, fairly, hundreds (of), in a way, kind of, kinda, loads (of), lots (of), many, maybe, more or less, nearly, not more than, occasionally, often, or anybody (like that), or anything, or so, or somebody, or something, or what ..., or whatever, or wherever, perhaps, probably, quite, rather, round, several, some, somehow, sometimes, somewhere, sort of, stuff, thousands (of).

The quantitative analysis involved comparison of frequencies in order to establish those VMs, whose use significantly differs in the two corpora in terms of frequencies. For this purpose the Log Likelihood test was used. The test shows the likeliness of such a difference occurring merely due to chance. If the Log-Likelihood value (further referred as LL) is lower than 3.84, it means that the difference is not statistically significant. However, if it is from 3.84 to 6.62 the probability that it is due to chance (as p) is lower than 0.05 (which is 5 % level). The online LL calculator was used to obtain the statistics which is available from <http://ucrel.lancs.ac.uk/llwizard.html>. The plus or minus sign preceding LL shows if a particular phrase is underused or overused in the selected corpus. The results of LL test will be presented in the first part of the paper.

For the qualitative analysis the authoritative Oxford Dictionary of English (OED 2003) was consulted for the meanings of the VMs. The qualitative analysis of VMs was based on the classification of VMs in Channell (1994), Jucker et al. (2003), Quirk et al. (1985) and Biber et al. (1999) whose classifications were used to design a specific system of the main VMs used in both corpora (see Fig. 5 for the alphabetical list of these categories).

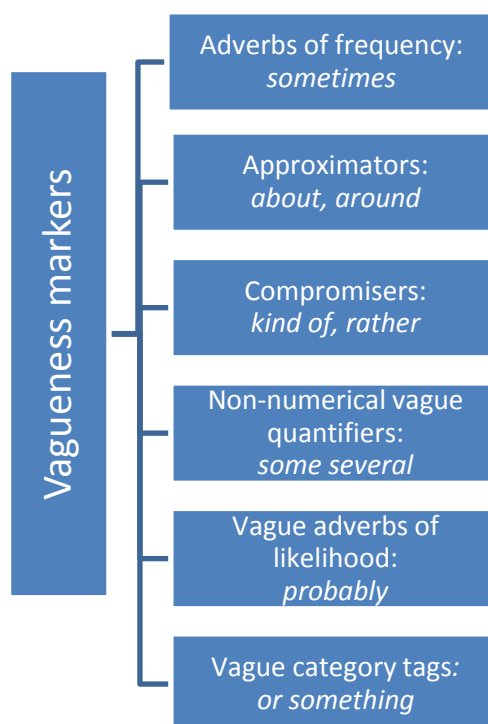


Fig 5. The classification of the main VMs used in LOCNEC and LINDSEI-LITH

It was noticed that compromisers and tags occur either in all three or at least two groups: underused by NNSs, overused by NNSs and used with a difference which is not significant. In addition, after conducting a pilot study some interesting differences in distribution of the

expressions, compromisers and tags in particular, were found. Thus, these two groups of VMs were chosen for the qualitative analysis, which is twofold: structural and functional.

For the qualitative analysis, firstly, it was tested whether the occurrences from LOCNEC realise uses which are described in Channell (1994) and grammars (Quirk et al. 1985 and Biber et al. 1999). Then, the other main tendencies in the usage of VL, for example, the usage with other parts of speech, and the functions of the whole sentence containing the VM in question, among NSs were identified. Finally, the tendencies in usage of VL by NNSs were compared to the ones by NSs. However, it is important to mention that the study remains quite subjective, as a number of occurrences could be classified and analysed in several different ways. Moreover, some of the VMs have more than one function, thus it is not easy to distinguish which is the main one and should be used in classifying the data.

Finally, the general conclusions from the results of both quantitative and qualitative parts of the paper will be drawn and the attempts will be made to answer the question ‘Why do NNS sound unnatural in their speech?’ regarding the issue of VL.

3. Results and Discussion

The quantitative analysis of 52 VMs will be presented in 3.1., whereas the qualitative analysis of nine VMs will be discussed in detail in the qualitative part of the paper (section 3.2.).

3.1. Quantitative analysis

Table 3 presents 10 most common VMs in both corpora. As it can be noticed, the most frequently used VM in LOCNEC is *sort of*, while in LINDSEI-LITH it is *maybe*. *Sort of* does not appear among the most common VMs in learners' corpus, whereas *maybe* is the last in the list showing top 10 in LOCNEC. Although a number of VMs appear on both lists, *about*, *(a) few*, and *stuff* are most common only in LOCNEC, while *many*, *a little (bit)* and *sometimes* are only among the ten top VMs in LINDSEI-LITH.

Table 3. Top 10 VMs in LOCNEC and LINDSEI-LITH

| | LOCNEC | Norm.fr. | LINDSEI-LITH | Norm.fr. |
|-----|----------|----------|--------------|----------|
| 1. | sort of | 358 | maybe | 323 |
| 2. | quite | 268 | some | 288 |
| 3. | some | 149 | quite | 163 |
| 4. | about | 123 | probably | 120 |
| 5. | a lot of | 89 | a lot of | 100 |
| 6. | probably | 70 | many | 84 |
| 7. | (a) few | 62 | kind of | 83 |
| 8. | kind of | 60 | a little | 82 |
| 9. | stuff | 59 | sometimes | 63 |
| 10. | maybe | 59 | a little bit | 62 |

All 52 VMs were checked applying Log-Likelihood test in order to find out whether the difference in frequency of the same VMs in different corpora is significant or not. See Table 4 for the overall frequency of VMs in LINDSEI-LITH in comparison to LOCNEC (see Appendix 2 for raw and normalised frequencies):

Table 4. The overall frequency of VMs in LINDSEI-LITH in comparison to LOCNEC

| | The number of VMs |
|---------------------------|-------------------|
| Underused | 19 |
| Overused | 12 |
| No significant difference | 21 |
| In total | 52 |

As the table shows, frequencies of 21 out of 52 VMs were not significantly different in the two corpora. However, almost the same number of VMs was underused. Finally, a number of VMs, 12 out of 52, are overused by NNSs.

The following sections (3.1.1., 3.1.2. and 3.1.3.) will report the main tendencies that were shown by Log-Likelihood test, regarding the underused, overused VMs by NNSs and those that are used with insignificant differences.

3.1.1. Underused vagueness markers in LINDSEI-LITH

19 VMs are significantly underused by NNSs (see Appendix 3 for the list of these VMs with p values). Table 5 shows that *sort of*, *about* and *and things* are found with the highest LL values: regarding *sort of* differences between the corpora are statistically significant at $p < 0.0001$, LL 435.17, while regarding *about* at $p < 0.0001$, LL 80.79 and *and things* at $p < 0.0001$, LL 66.52.

Table 5. The LL values for the underused VMs by NNSs

| | Vagueness marker | LL | | Vagueness marker | LL |
|-----|------------------|--------|-----|------------------|---------|
| 1. | et cetera | -3.85 | 11. | rather | -13.41 |
| 2. | occasionally | -4.81 | 12. | fairly | -14.42 |
| 3. | nearly | -5.77 | 13. | loads (of) | -14.44 |
| 4. | anywhere | -6.63 | 14. | and stuff | -17.61 |
| 5. | and everything | -6.88 | 15. | quite | -28.39 |
| 6. | (a) few | -8.91 | 16. | a bit of | -36.73 |
| 7. | or anything | -9.14 | 17. | and things | -66.52 |
| 8. | a couple of | -9.8 | 18. | about | -80.79 |
| 9. | or whatever | -10.45 | 19. | sort of | -435.17 |
| 10. | stuff | -11.39 | | | |

Interestingly, four VMs, *fairly*, *nearly*, *occasionally*, *et cetera*, have only been found in LOCNEC and none of them appears in LINDSEI-LITH. Three significantly underused VMs, *sort of*, *quite*, *rather*, belong to the group of compromisers and five, *and things*, *and stuff*, *or whatever*, *or anything*, *and everything*, to tags, and thus will be included into the qualitative analysis.

3.1.2. Overused vagueness markers in LINDSEI-LITH

Table 6 lists all 12 significantly overused VMs by learners (see Appendix 4 for the p values). The highest LL values show that *maybe*, *some*, *and so on* are the most overused VMs.

Differences of the frequency of *maybe* between the corpora are statistically significant at $p < 0.0001$, LL 238.44, *some* at $p < 0.0001$, LL 51.54 and *and so on* at $p < 0.0001$, 42.96.

Table 6. The LL values for the overused VMs by NNSs

| | Vagueness marker | LL | | Vagueness marker | LL |
|----|------------------|--------|-----|------------------|---------|
| 1. | kind of | +4.44 | 7. | Somehow | +22.38 |
| 2. | several | +6.73 | 8. | a little bit | +34.7 |
| 3. | sometimes | +14.49 | 9. | Many | +40.37 |
| 4. | probably | +15.18 | 10. | and so on | +42.96 |
| 5. | kinda | +17.03 | 11. | Some | +51.54 |
| 6. | a little | +18.69 | 12. | Maybe | +238.44 |

And so on, which is a tag, and also the compromiser *kind of*, which is also overused in LINDSEI-LITH, will be discussed in the qualitative analysis.

3.1.3. Frequencies of vagueness markers with insignificant statistical differences

Table 7 shows the list of 21 VMs whose frequencies were not significantly different in the two corpora (see Appendix 5 for the p values).

Table 7. The LL values for the VMs that occur without a significant difference regarding the frequency

| | Vagueness marker | LL | | Vagueness marker | LL |
|-----|------------------------|-------|-----|------------------|--------|
| 1. | perhaps | -2.62 | 12. | often | -0.1 |
| 2. | in a way | -1.81 | 13. | more or less | + 0.01 |
| 3. | round | -0.96 | 14. | approximately | +0.06 |
| 4. | not more than | -0.96 | 15. | or something | +0.18 |
| 5. | or somebody | -0.96 | 16. | or so | +0.58 |
| 6. | or anybody (like that) | -0.96 | 17. | a lot of | +0.7 |
| 7. | or wherever | -0.96 | 18. | thousands (of) | +1 |
| 8. | or what ... | -0.77 | 19. | somewhere | +1.86 |
| 9. | lots (of) | -0.57 | 20. | around | +1.93 |
| 10. | hundreds (of) | -0.31 | 21. | almost | +2.62 |
| 11. | and all that | -0.29 | | | |

Although the LL test shows that 12 out of these 20 VMs are underused by learners and 8 are overused, the probability that differences between the two corpora occur due to the chance is very high. Moreover, there are even six tags in this group; however, the usage of most of them

is very low in both corpora. Due to this reason, and also because the analysis of the significantly underused and overused VMs is likely to give more reliable results, these VMs will not be included into further analysis.

3.2. Qualitative analysis

The results of the qualitative analysis of vague coordination tags (hereafter tags) and compromisers will be presented in the following sections. In the beginning of each section the grammatical and functional classifications mainly by Quirk et al. (1985), Biber et al. (1999) and Channell (1994) will be presented, which will be followed by the discussion of general tendencies in LOCNEC and the structural and functional analysis of underused and overused VMs in LINDSEI-LITH. Due to lower frequencies of tags when compared to compromisers, firstly, six tags will be overviewed (section 3.2.1.), secondly, three most often occurring compromisers in learners' corpus will be analysed in depth (sections 3.2.2., 3.2.3., 3.2.4.).

3.2.1. Vague coordination tags

In this section the most common grammatical and functional characteristics by Channell (1994), Biber et al. (1999) and Quirk et al. (1985) will be presented and, then the general tendencies, regarding all six tags will be overviewed.

According to Channell (1994: 119), reference to categories is one of the major areas where vagueness is used in English. The structure that is often used for this purpose, usually starting with *and/or* followed by some other word(s), is called a tag, e.g. *and things, or something*. Sometimes tags have a function of performance 'fillers' that just give a speaker and hearer additional time for processing; however, in a number of cases removing tags results in changing the meaning of the sentence, and in order to retain the same meaning additional paraphrase is needed (ibid. pp. 120-121) as in (3). Thus, Dines (1980:19 in Channell, 121) refers to *tags* as "[salient elements] which qualify the content of the utterance".

(3) *You can remember four lots of four fairly easily say in the form of dates **or something**, 1972 **or something like that** (original sentence)*

(4) *You can remember four lots of four fairly easily say in the form of dates, 1972 (tags are omitted) (Channell 1994:121)*

In order to recover the same meaning as in (3), the sentence in (4) has to be paraphrased, e.g.: 'say in the form of dates, like 1972'. Therefore, the main function of *vague tags* seems to be to indicate that the preceding element should be understood in a more general way. Channell concludes that 'the whole expression [the tag and the element(s) preceding it] directs the hearer to access a set, of which the given item is a member whose characteristics will enable

the hearer to identify the set' (1994: 122), and thus the tags can be called vague category identifiers. The results of Channell's test with NSs revealed that NSs '[interpret] the tags as instructions to access categories, and (...) that the characteristics of the categories they fixed on are determined not only by the given exemplar, but also by linguistic context, presumed situation, and pragmatic knowledge' (ibid. p. 131). Moreover, they can also be referred to as vague additives, because they are added to the statements that would otherwise be precise (ibid. p. 18). Thus, semantically, they are subordinate, and instead of having independent reference, these elements just indicate that the expression followed by the conjunction and the tag is not precise or exhaustive (Biber et al. 1999: 116).

A number of tags contain the phrase *like that*. Channell proved that there is no difference how the tags with this phrase and without are understood, and thus the tag without it could be treated as ellipted version of the full one. Furthermore, *or something* and *or anything* follow a rule of special marking in non-assertive contexts, and thus *or something* occurs in assertive contexts, while *or anything* in negative contexts (ibid. p. 132; Quirk 1985: 782), e.g.:

(5) *There was no coffee room or anything.*

(6) *One of the secretaries was saying there was a film or something.* (Channell 1994: 132)

Channell's research showed that the most common word classes followed by *tags* are nouns (most often countable singular or countable plural and sometimes uncountable) and verbs. It is important to mention that tags occurring with nouns tend not to be subjects. In addition, the exemplar, which is the element(s) preceding the tag, provides some new information. Finally, the tags can act as list completers, occurring in, e.g. the pattern similar to '*noun+noun+tag*'. Prepositions are also sometimes followed by *tags*, and there are only few tags used with adverbs and adjectives (ibid. pp. 132-135).

The participants of Channell's test admitted that there are two conditions on successful vague tagging: '[t]here must exist other members of the set which the tag can stand for' and '[t]he hearer must make the relevant category boundaries at least similar to the speaker' (ibid. p. 141). Thus, the sentence 'It's a boy or a girl or something' would sound odd. The second condition refers to being uninformative and failing to give expected information, therefore also violating the maxim of Quantity. To add more, Biber et al. (1999: 115-116) claim that elements coordinated by tags are related in form and meaning, thus usually but not always having connection with each other and having the same relationship to surrounding structures. The elements preceding and following the conjunction cannot be reordered.

Biber et al. (1999: 116) show that the most common coordination tag in conversation is *or something*, the second most common is *and everything*, the third most common is *and things (like that)*, and the fourth one is *and stuff (like that)*. Table 8, displaying normalised frequencies of vague tags starting with *and/or*, shows that these four tags are also most commonly used by NSs with a slight difference in order: *and things* and *or something* are most often used tags, while *and everything* is used less often. *And stuff* is in the fourth place.

Table 8. Vagueness tags from most frequent to least frequent

| | Tags in LOCNEC | Normalised fr. | | Tags in LINDSEI-LITH | Normalised fr. |
|-----|------------------------|-----------------------|-----|-----------------------------|-----------------------|
| 1. | and things | 58 | 1. | or something | 58 |
| 2. | or something | 53 | 2. | and so on | 32 |
| 3. | and everything | 33 | 3. | and everything | 16 |
| 4. | and stuff | 35 | 4. | and stuff | 9 |
| 5. | or anything | 18 | 5. | or so | 6 |
| 6. | or whatever | 14 | 6. | or anything | 4 |
| 7. | and all that | 5 | 7. | and all that | 3 |
| 8. | or so | 3 | 8. | and things | 2 |
| 9. | or what ... | 3 | 9. | or whatever | 2 |
| 10. | and so on | 1 | 10. | or what ... | 1 |
| 11. | or somebody | 1 | 11. | or somebody | 0 |
| 12. | or anybody (like that) | 1 | 12. | or anybody (like that) | 0 |
| 13. | or wherever | 1 | 13. | or wherever | 0 |

The use of *or something* is similar in both corpora: normalised frequency (further *norm. fr.*) being 53 in LOCNEC and 58 in LINDSEI-LITH. However, the most often used tag by NSs, *and things*, is one of the least often used by NNSs: *norm. fr.* in LOCNEC is 58, while in LINDSEI-LITH it is only 2. The tag *and everything* is quite common in both corpora. Interestingly, *and stuff* is the fourth most common vague tag in native speakers' and learners' language; however, the normalised frequency shows that it is used almost four times more often by NSs. Moreover, *and so on*, which is more common in the written register (Biber et al. 1999: 116) is very often used by NNSs.

As it can be seen from the LL values in Table 9, *and things* is a tag that is most underused by NNSs:

Table 9. LL values of underused and overused tags

| | LL |
|----------------|-----------|
| and so on | +42.96 |
| and everything | -6.88 |
| or anything | -9.14 |
| or whatever | -10.45 |
| and stuff | -17.61 |
| and things | -66.52 |

The difference in frequencies in both corpora is also very huge concerning the other three tags: *and stuff*, *or whatever* and *or anything*. The difference in usage of *and everything* is not that huge if compared to the previously mentioned tags but still highly significant. In addition, there is only one tag that is overused by Lithuanian learners of English, but the LL value is very high. There are only two VMs with higher LL: *some* (LL 51.54) and *maybe* (LL 238.44). However, due to the limited time for the research as well as scope of the MA paper, these two will not be discussed in the paper.

The underused (sections 3.2.1.1., 3.2.1.2.) and overused (sections 3.2.1.3., 3.2.1.4.) tags by NNSs will be discussed in more detail in the following chapters.

3.2.1.1. Structural analysis of underused tags in LINDSEI-LITH

The next table presents the use of different word classes with tags, displaying the word class of the exemplar it follows and modifies.

Table 10. Word class used with tags by NSs and NNSs, %

| | nouns | | verbs | | adjectives | | other | |
|----------------|--------------|-------------|--------------|-------------|-------------------|-------------|--------------|-------------|
| | NSs | NNSs | NSs | NNSs | NSs | NNSs | NSs | NNSs |
| And everything | 72.9 | 85.7 | 10.4 | 7.1 | 6.3 | 7.1 | 8.3 | 0 |
| And stuff | 54.9 | 62.5 | 35.3 | 12.5 | 5.9 | 0 | 3.9 | 25 |
| And things | 77.6 | 100 | 12.9 | 0 | 1.2 | 0 | 8.2 | 0 |
| Or anything | 61.5 | 25 | 23.1 | 25 | 15.4 | 25 | 0 | 25 |
| Or whatever | 81.0 | 100 | 4.8 | 0 | 4.8 | 0 | 9.5 | 0 |

The results show that the majority of tags in LOCNEC (from 54.9% to 81 % of all occurrences with the particular tag) follow and modify nouns. In LINDSEI-LITH all occurrences of *and things* and *or whatever* contain nouns. The majority of *and stuff* (62.5%) and *and everything* (85.7%) also follow the same pattern, while only 25% of all the

occurrences of *or anything* contain a noun. However, there are only 4 occurrences of this tag in LINDSEI-LITH, thus there is one tag following a noun, one following a verb, one following an adjective, and one modifying a clause. Therefore, due to such low usage of the phrase by NNSs, it is difficult to draw any conclusions.

The analysis of the tags in LOCNEC displays a clear preference to countable nouns. Although the frequency of these tags is low in LINDSEI-LITH, it can be noticed that the usage of some tags in this corpus shows preference to uncountable nouns, e.g. *or stuff*, *or whatever*, while the others, *and things*, *or anything*, to countable nouns. Finally, half of the occurrences of *and everything* contain countable nouns, the other half uncountable. See Tables 11, 12:

Table 11. Countable and uncountable nouns, LOCNEC

| | nouns in total | Count. | % | Uncount. | % |
|----------------|----------------|--------|-------|----------|------|
| and things | 66 | 58 | 87.9* | 12 | 18.2 |
| and stuff | 28 | 19 | 67.9 | 13 | 46.4 |
| or whatever | 17 | 10 | 58.8 | 7 | 41.2 |
| or anything | 16 | 15 | 93.8 | 1 | 6.3 |
| and everything | 35 | 28 | 80 | 9 | 25.7 |

*In some occurrences there are countable and uncountable nouns

Table 12. Countable and uncountable nouns, LINDSEI-LITH

| | nouns in total | Count. | % | Uncount. | % |
|----------------|----------------|--------|-----|----------|-----|
| and things | 2 | 2 | 100 | 1 | 50 |
| and stuff | 5 | 2 | 40 | 3 | 60 |
| or whatever | 2 | 0 | 0 | 2 | 100 |
| or anything | 1 | 1 | 100 | 0 | 0 |
| and everything | 12 | 6 | 50 | 6 | 50 |

All these tags are also sometimes used after verbs and adjectives by NSs. For example, even 35.3 % of all the occurrences of *and stuff* contain verbs and 23.1% of *or anything* follow this word class as well. In most cases, NNSs use tags with verbs and adjectives less often than NSs. In addition, the section ‘other’ in Table 9 indicates that the tags were also found in some other constructions, e.g. in LOCNEC 7 occurrences, which comprise 8.2% of all sentences with *and things*, were found changing the meaning of the whole clause instead of only a phrase (compare examples (7) and (8)). Similarly, 2 occurrences (3.9%) of *and stuff* in LOCNEC and 2 occurrences (25%) in LINDSEI-LITH are also added to the clause.

(7) [*there's a lot of parties* <\B>

 and things but whereas here we go out a lot more I mean .. we don't do anything
<XX> I mean Furness you know like a corridor of about twelve <\B> (E12)

(8) and er she says <X> probably got a gun behind him and .. and I was like oh
<laughs> <\B>

 it's it's just something I couldn't comprehend **and things** like that <\B> (E17)

In (7) *and things* is attached to the noun 'parties', while in (8) it is attached not only to 'couldn't comprehend', but to the whole clause 'it's just something I couldn't comprehend'. In addition, *or whatever* in two sentences (9.5%) found in LOCNEC modify numbers as in (9). *And everything* in four sentences produced by NSs is also added to the clause, comprising 8.3% with 'other' usage (see example (10)).

(9) in Cambridge <X> we met up with other Film Societies and a lot of them were were
sort of like . baby societies they'd [only just started and they only had sort of **thirty**
members **or whatever** <\B> (E15)

(10) yeah you had each time there was a violent act I had two forms to fill in .. so it was
a lot of work <\B>

 [but it was it was not a bad job you know I could do it at home <\B>

 and everything <\B> (E30)

Although in (9) the tag follows a noun, but it modifies a number, thus it was included into the group 'other'. In (10) it is clear that the exemplar is the whole clause 'I could do it home', not just the word 'home'. Finally, one occurrence with *or anything* (25%) is also found after the clause in LINDSEI-LITH, while there are no occurrences of this tag in LOCNEC belonging to the section 'other'. However, overall, the table also shows that such 'other' constructions are more often used by NSs rather than NNSs.

Four tags, *and things*, *and stuff*, *or anything*, and *and everything*, were also found in the full constructions containing 'like that'. *And things like that* and *and stuff like that* are very common patterns among NSs. The former occurs in 33 occurrences out of 88, and the latter in 17 out of 51. They are also common among NNSs: *like that* is present in both sentences with *and things*, and *and stuff* in four occurrences out of eight. In LOCNEC the tags are found in the sequences *or anything like that* and *and everything like that* just few times, while these sequences have not been found in LINDSEI-LITH.

The only underused tag which is not followed by 'like that' in neither of the corpora is *or whatever*. However, few instances were found in both corpora when this tag was used with

more VL in the same sentence, e.g. *sort of, how you call it, I don't know*. Similarly, *kind of* and *sort of* are found together with *or anything* in LOCNEC, e.g.:

(11) *and language acquisition* <\B>

 at the moment .. simply because I'm interested in language acquisition it's not as a kind of career or anything <\B> (E19)

(12) *not supposed to be in a sort of safari or anything . just thinking* <X> *if you were in a car . you know (...)* (E51)

However, there were no such instances of this tag in LINDSEI-LITH.

3.2.1.2. Functional analysis of underused tags in LINDSEI-LITH

The underused tags perform two main functions that are: to indicate that the preceding element or elements should not be understood as precise or that it or they are not exhaustive (the function of a list completer). In addition, some of the tags, e.g. *and things, or anything, and everything*, in few occurrences in LOCNEC function as performance fillers as in (13) and (14):

(13) *but then you: you spend a fortune photocopying . [and I've done that* <\B>

 <XX> *oh* <\B>

 [*yeah it is with I know the card it's six pence* <\B>

 rather than t= ten pence but even so you know you're doing a lot of pages and things <\B>

 you get a card . let me show it to you <\B> (E11)

(14) *travelling most of the time well .. it' d be difficult to like come to the socials at night*

<\B>

 <X> *I'll have to think of going home* <X> *I wouldn't be able to drink or anything so*

<\B>

 and so I .. I wouldn't be .. able to enjoy it as much <XX> <\B> (E02)

The context of (13) suggests that the interviewee is speaking about photocopying and complaining about the amount of money he or she needs to pay because of the huge number of pages. The speaker already mentioned 'doing a lot of pages', and thus it is not clear what else the set including the exemplar contains. Similarly, in (14) the speaker by saying 'I wouldn't be able to drink or anything' is perhaps just conveying the message that he or she 'would not be able to consume alcohol', and thus the members of the set are not clear, unless we go extreme, and claim that the interviewee is having drugs in mind as well. Thus, even if these tags were removed, it would not change the meaning of the sentence. Table 13 shows the frequency of tags functioning as performance fillers.

Table 13. Performance fillers, raw frequencies

| | NSs | | NNSs | |
|----------------|------------|-------|------------|-------|
| | Perform.f. | Total | Perform.f. | Total |
| and everything | 2 | 48 | 0 | 14 |
| and stuff | 0 | 51 | 0 | 8 |
| and things | 1 | 85 | 0 | 2 |
| or anything | 2 | 26 | 0 | 4 |
| or whatever | 0 | 21 | 0 | 2 |

Thus, the usage of tags with such function is low among NSs and is has not been found among NNSs regarding all the underused tags.

Table 14 shows the frequencies of the tags functioning as indicators of imprecision of the preceding element(s) and list completers.

Table 14. Frequencies of tags with particular functions, raw frequencies, %

| Tags | Indicator of imprecision (function 1) | | | | List completer (function 2) | | | |
|----------------|---------------------------------------|------|------|------|-----------------------------|------|------|------|
| | NSs | % | NNSs | % | NSs | % | NNSs | % |
| and everything | 16 | 33.3 | 4 | 28.6 | 30 | 62.5 | 9 | 64.3 |
| and stuff | 34 | 66.7 | 4 | 50 | 25 | 49.0 | 4 | 50 |
| and things | 49 | 57.6 | 1 | 50 | 61 | 71.8 | 1 | 50 |
| or anything | 19 | 73.1 | 3 | 75 | 5 | 19.2 | 1 | 25 |
| or whatever | 21 | 100 | 2 | 100 | 0 | 0 | 0 | 0 |

All five tags appear with the function of a list completer in a number of occurrences in LOCNEC, while NNSs use only three of them with this function: *and things*, *and stuff*, and *everything*. A tag is considered to possess this function if it follows more than one word of the same word class or occur in the same structure (see example (15)).

(15) [mhm .. yeah yeah .. yeah .. well the Scout and Guide Club here is more it's more scout orientated we do more **camping and hiking ..and [things like that** <\B> (E13)

For example, in (15) *and things like that* stands for similar activities to camping and hiking. Thus, the VM indicates the incompleteness of the list. Moreover, the table shows that *or whatever* is used only function 1 by NSs as in (16) as well as NNSs. Both groups of speakers also use *or anything* more often with this function (see example (17)), while *and everything* is much more often used with function 2 as in (18).

- (16) *at the beginning of the term . and erm .. she was upset over that and she was still carrying a lot a lot of er .. what you would call you know*
 emotional baggage or whatever .(...) (E07)
- (17) <overlap /> *yes before before the movie*
 I I like I liked the book very much it was well it was very long about six hundred pages something like that (erm) but it wasn't intense or anything it was more like philosophical things and I guess it's quite hard to translate it to the movie (LT010)
- (18) (...) *so really that was probably it and because it's not really (eh) of course the architecture the culture and everything was really impressive in Belgium but by by far the be= best thing about the country is the people so yeah* (LT011)

The phrase ‘what you would call you know’ in (16) clearly indicates that the speaker is not sure what words to use, and the tag *or whatever* proves that and indicates that the phrase ‘emotional baggage’ is not precise. In (17) *or anything* performs the same function, because for the speaker the adjective ‘intense’ perhaps does not convey the exact meaning that he or she has in mind.

While previous examples illustrated the similarities, the results concerning *and things* and *and stuff* reveal some interesting differences. Firstly, interviews with NSs show preference for function 2 with the former, whereas they use the latter more often with function 1. To contrast, 50% of the instances of *and things* in LINDSEI-LITH are used to emphasise the imprecision, and the other half to show that the list is not exhaustive. The same situation is with *and stuff*: half of the phrases are used with function 1 and the other half with function 2. Consider the following examples:

- (19) *yeah <laughs> [I'm supposed to read er . the Guardian and things like that*
 (E06)
- (20) (...) *one day we to bought some coffee and went to a park and we had a great time because people all around were having lunch and eating salad and stuff like that so it was great I loved it* LT031

In (19) *and things like that* stands for similar newspapers, indicating that a list is much longer, while in (20) it seems that the speaker wants to convey the message that people were not necessarily having lunch there, and thus ‘having lunch and eating salad’ should not be understood as precise. Thus, it could be concluded that NSs use the majority of underused tags more often in order to indicate the imprecision of the preceding elements, while this is true only for few tags in NNS language.

Another interesting aspect that should be mentioned is that quite a huge number of instances of *and things*, even 27 out of 85, used by NSs can be analysed in both ways, e.g.:

(21) mm it is .. once you've been living here a few years you know everybody in Lancaster it's it's a nightmare you meet [somebody who <\B>

 yeah well you you meet somebody who you think is completely new and you think oh you know I've met someone new and exciting or I don't know .. and it turns out they know all your friends from school **and things** <laughs> <\B> (E16)

(22) yeah we haven't got enough money so <laughs> <\B>

 yeah I've been on day trips with school **and things** like that to France but [that's it <\B>

 well it doesn't really count <laughs> <\B> (E29)

Thus, in (21) the purpose of using the tag could be to indicate that the list is incomplete, and the speaker emphasises that the person you thought to be new knows not only your friends from school, but your friends from somewhere else, e.g. your work, or knows your family or relatives. On the other hand, it could also be claimed that the exemplar 'all your friends from school' should not be taken word for word but meaning something more general as 'they know much more about you than you thought they do'. Similarly, in the second example 'day trips with school' could be understood just as an example of types of short trips a person had or it could be interpreted as 'some short trip with school, but not necessarily lasting a day'. Such occurrences are included into both sections of table 13 (*function 1* and *function 2*). The other tags used by NSs and all the tags used by NNSs are interpreted clearly having one or the other function.

It was also noticed that learners very often use *and everything*, when clearly giving some examples, e.g.:

(23) but what I like it here is that it's more . the city is more alive if I could say like this the= there is always some action (er) in the city center for example and **and everything** and . (er) and when I come back now in Klaipeda I don't find really much what to do there (...) (LT004)

The speaker even uses the phrase 'for example' in (23). In LOCNEC this function was not that visible.

3.2.1.3. Structural analysis of overused tags in LINDSEI-LITH

The following table presents the use of different word classes with the only tag that is overused by learners.

Table 15. Word classes with and so on, raw frequencies

| VM | nouns | | verbs | | adjectives | | other | |
|-----------|-------|------|-------|------|------------|------|-------|------|
| | NSs | NNSs | NSs | NNSs | NSs | NNSs | NSs | NNSs |
| and so on | 0 | 19 | 2 | 6 | 0 | 0 | 0 | 4 |

The table shows that NSs used both occurrences after verbs, while among NNSs the most common word class preceding *and so on* is noun (16 countable nouns, eight uncountable). Some instances also contain verbs, but they constitute only 20.7% of all the occurrences of *and so on*. Four tags in LINDSEI-LITH also follow a clause, and thus in the table are included into the section ‘other’.

3.2.1.4. Functional analysis of overused tags in LINDSEI-LITH

As regards functions, both tags in LOCNEC function as performance fillers, but there are no occurrences with this function in LINDSEI-LITH (see table 16).

Table 16. Functions of and so on, raw frequencies, %

| VM | Performance fillers | | | | Indicator of imprecision | | | | List completers | | | |
|-----------|---------------------|-----|------|---|--------------------------|---|------|------|-----------------|---|------|------|
| | NSs | % | NNSs | % | NSs | % | NNSs | % | NSs | % | NNSs | % |
| And so on | 2 | 100 | 0 | 0 | 0 | 0 | 5 | 17.2 | 0 | 0 | 24 | 82.8 |

Moreover, over 82% of all the occurrences of *and so on* in LINDSEI-LITH show that the exemplar is not exhaustive, and thus there is something to add. To add more, in 19 out of 29 sentences produced by NNSs the tags act as list completers, and similarly as with *and everything*, the speakers are giving some examples before adding *and so on*. This function has not been found in LOCNEC. Consider the following examples:

(24) erm so I've been living . in town I've rent a house in town I've been living there for about three years so it seemed a bit pointless to sort of <X> move onto campus <\B>
 and move back off again next year **and so on** <\B> (E16)

(25) but erm .. it was just different from all the[i:] other films I'd ever seen <X> that style because you never ever saw any of the murders ever happening you just sort of saw the victims afterwards . and <\B>
 yeah once once yeah they'd been killed **and so on** and er and you didn't actually get any of the . the[i:] actual acts happening <\B> (E28)

(26) (...) I don't know (erm) I think it's nature . (mm) well my first time in Spain was . was like an excursion so I I visited Barcelona and I visited its churches its its (eh) its famous

places monuments and so on and the other times I was like I spent there the whole summer so it was more like living there (...) (LT025)

(27) (...) *I liked that the s= the streets were so (em) nice and tidy and (er) I liked also that there was a lot of nature like trees and water and parks and so on (...)* (LT048)

In (24) and (25), the other members of the sets are not clear, and it seems that the speaker simply refers to moving to a house from campus in the first example, and to actual killing in the second example. Thus, the tags could be removed from these sentences without the change of meaning. The exemplar is a list completer in (26), and thus it shows that previously mentioned elements are not exhaustive, while in the last example, the speaker is giving examples.

Thus, *and so on* is very differently used by these two groups of speakers. This tag is much more often used in written discourse, and when used in conversation by NSs, it does not convey any message. On the contrary, it seems that NNSs use the tag in a very similar way as they would use it in writing.

3.2.2. Compromiser *quite*

This chapter will provide the overview of the grammatical classifications and characteristics of the compromiser *quite* by Quirk et al. (1958) and Biber et al. (1999), which is underused by NNSs, LL -28.39.

First of all, OED provides two main meanings of this adverb: 1) to the utmost or most absolute extent or degree; absolutely; completely (US: very; really), 2) to a certain or fairly significant extent or degree (2003: 1445). These meanings correlate with the classification and functions of *quite* found in grammars. Although most often *quite* seems to be functioning as a compromiser, and thus it is a downtoner, according to Quirk et al. (1985), it may also function as an amplifier, as it was shown in data provided in Fig. 4.

Before going deeper into the distinction between amplifiers and downtoners, the gradability of a word should be discussed. Verbs that are referred as gradable usually denote actions that can happen more or less completely or fully, e.g. actions that are related to feelings or emotions like *love, like*. Gradable adjectives are those that allow comparison and degree modification, e.g. they take comparative and superlative forms and follow adverbs of degree like *very* and *extremely*. On the other hand, adjectives that serve as classifiers which place a noun's referent in a category in relation to other referents, and thus delimit or restrict it, e.g. criminal law, medical students, urban district, or that have an identifying or intensifying meaning (e.g. the identical car, utter nonsense) are usually non-gradable (Biber et al. 1999: 64; 506-508).

However, some non-gradable adjectives can be modified by emphatic adverbs, for instance, *quite motionless*, *really tremendous*, *absolutely continuous* (ibid. p.521).

As regards *quite*, if it co-occurs with a gradable adjective or adverb, e.g. *quite good*, it is a *downtoner*, and usually means ‘a little, moderately but not very’ or it has a similar meaning to ‘rather’ or ‘fairly’, while if used with a non-gradable adjective or adverb, e.g. *quite right*, it has the function of an *amplifier*, thus usually meaning ‘very, totally or completely’ (CAED)². However, despite this rule, in AmE even though *quite* precedes a gradable adjective/adverb, it can still mean *very* (Quirk et al. 1958: 446). Moreover, some adjectives can be interpreted as gradable or as non-gradable as *confident* in (28), where *quite confident* refers either to moderate or complete confidence (Biber et al. 1999: 556). Thus, in order to analyse the functions of these *intensifiers* the context and intonation have to be taken into consideration. However, for the purpose of this analysis the latter is impossible as the corpora lack prosodic marking.

(28) *I was quite confident that it would stay in.* (ibid.)

In order to find out whether the modifying adverb is in fact an *amplifier*, a semantic test, which includes checking whether the word in question can be contrasted in alternative negation with the vague phrase *to some extent*, has to be applied. According to this test, the adverb in (29) is an amplifier, but the one in (30) is not:

(29) *He didn't ignore my request completely, but he did it to some extent.*

(30) **He didn't really ignore my request, but he did ignore it to some extent.* (Quirk et al. 1958: 590-591)

Moreover, being classified as an amplifier, according to Quirk et al. (ibid.), *quite* falls under the category of maximizers. They also add that only maximizers are found with negatives as in (31):

(31) *She didn't quite approve.*

Furthermore, being classified as a downtoner, *quite*, may function not only as a compromiser, but also as a diminisher. The former are usually stressed only lightly, while the latter are heavily stressed or even made nuclear. Consider the following examples:

(32) *The book is quite GOOD.* [*compromiser or amplifier*]

(33) *It seems that they quite LIKE her.* [*compromiser*]

(34) *The book is QUITE GOOD.* [*diminisher*]

² See <http://dictionary.cambridge.org/grammar/british-grammar/quite> for more information on this rule

(35) *It seems that they QUITE LIKE her. [diminisher]* (Quirk et al. 1958: 599)

As it can be seen from the examples, without hearing the intonation and without the knowledge which words were stressed, it is very difficult to distinguish compromisers from diminishers; however, the context can help to make a distinction between these two.

The structural analysis of *quite*, including syntactic patterns and lexical combinability, will be discussed in section 3.2.2.1., whereas the functional will be presented in 3.2.2.2.

3.2.2.1. Structural analysis of *quite*

It was found that most commonly used word class with *quite* is adjective: in over 50% of all the occurrences with *quite* this VM is followed by an adjective (see Table 17). In both corpora the second most common word class is noun. Interestingly, the occurrences with adjectives as well as the ones with nouns comprise the bigger part in learners' corpus than in LOCNEC. Thus, there are fewer syntactic patterns with verbs and adverbs in LINDSEI-LITH than in LOCNEC. Moreover, there are significantly fewer occurrences extracted from NNS corpus that fall under the group of 'Other', which means that the VM is used at the end of a sentence or unfinished phrase, introduces a clause or cannot be classified as the other occurrences for some other reason. Hence, it seems that learners just follow the most common patterns.

Table 17. Word class following *quite* in LOCNEC and LINDSEI-LITH

| | Nouns | Verbs | Adjectives | Adverbs | Other | Total |
|-----------------------------------|-------|-------|------------|---------|-------|-------|
| Percentage in LOCNEC | 18.6% | 5.1% | 51.3% | 13.3% | 11.3% | 100% |
| Percentage in LINDSEI-LITH | 26.5% | 4.1% | 56.5% | 8.8% | 4.1% | 100% |

Table 18 presents the most common syntactic patterns of *quite* and adjectives (adj) (see Appendix 6 for all the syntactic patterns with adj). The table shows that the most common pattern in both corpora is *be* in Present Simple (Pr. S.) followed by *quite*, which precedes the adjective. This construction is even more common in LINDSEI-LITH than LOCNEC, as in the former it comprises 27.9% of all the occurrences with *quite*, while in the latter 21.2%. The second most common pattern among NSs is *quite+adj*, occurring with another verb rather than *be* or without a verb at all. The frequency of this construction is almost the same as of the construction which is almost identical to the most common one, except that *be* is in Past Simple (Past S.) However, *be(Past S.)+quite+adj* is much more frequently used among NNSs than *quite+adj*. It could also be noted that *quite+as+adj* occurred few times in LOCNEC;

however this syntactic pattern has not been found in LINDSEI-LITH. Examples (36)-(38) illustrate the most common syntactic patterns:

(36) (...) *I am quite happy to be here and . well (eh) the main drawback is that very often I (em) . go to bed very late because you know we have to do a lot of work and and <laughs> since I like to sleep a lot so (LT002)*

(37) (erm) as he was painting the queen he painted her quite realistically and everyone around him **was quite shocked** so was the queen when she saw <laughs> the[i:] final not the final the draft of it so and in the second like this frame (...) (LT010)

(38) even subtitled ones **seem quite good** <begin_laughter> at times you know <end_laughter> as long as you can concentrate on them I think
 but er .. [that's the good thing about it (E22)

Table 18. Syntactic patterns with *adj.*, raw fr., %

| | Number of occurrences in LOCNEC | Percentage in LOCNEC | Number of occurrences in LINDSEI-LITH | Percentage in LINDSEI-LITH |
|-----------------------|---------------------------------|----------------------|---------------------------------------|----------------------------|
| be(Past S.)+quite+adj | 61 | 15.6 | 26 | 17.7 |
| be(Pr. S.)+quite+adj | 83 | 21.2 | 41 | 27.9 |
| quite+as+adj | 4 | 1.0 | 0 | 0.0 |
| quite+adj | 64 | 16.3 | 15 | 10.2 |

Quite is one of the most common adverbs that are treated as intensifiers that can premodify nouns and also precede the determiner in doing so (Biber et al. 1999: 450). Table 19 presents the most common syntactic patterns with *quite* and nouns (see appendix 7 for all the patterns) and shows that the same two patterns, *quite+article+adj+n* and *quite+a lot of+(article)+n*, are most often found in both corpora. Interestingly, both patterns account for the bigger part of all the occurrences with *quite* in LINDSEI-LITH. Although the difference is not that huge with *quite+article+adj+n*, it is much more visible with *quite+a lot of+(article)+n*: this syntactic pattern accounts almost for 5% of all the occurrences with *quite* in LINDSEI-LITH, but only 2.6% in LOCNEC. In addition, *quite+adj+n* accounts for 4.1% among NNSs, while the use of this pattern is insignificant among NNSs at all. Thus, with nouns learners are using

those syntactic patterns that are not that common in NS language. Consider the following examples for the most common usage:

(39) <laughs> erm . there was a number of things really erm ... one was this er .. **quite a practical reason** really this erm . the job market was <begin_laughter> so very <end_laughter> bad when I left school (E01)

(40) (...) this year there were **quite a lot of people** participating (erm) from various countries like (em) Italy Poland Russia (er) even Great Britain (...)(LT004)

(41) when I got my my first camera . **quite professional camera** yeah (LT056)

(42) yeah <X> you're gonna get as cl= claustrophobic I mean I was a bit apprehensive before I came I'd not be able to meet anybody you know I was . living **quite a way** but that's not been too much of a problem (E45)

Table 19. Syntactic patterns with nouns

| | Number of occurrences in LOCNEC | Percentage in LOCNEC | Number of occurrences in LINDSEI-LITH | Percentage in LINDSEI-LITH |
|----------------------------|---------------------------------|----------------------|---------------------------------------|----------------------------|
| quite+article+adj+n | 23 | 5.9 | 10 | 6.8 |
| quite+article+n | 9 | 2.3 | 1 | 0.7 |
| quite+a lot of+(article)+n | 10 | 2.6 | 7 | 4.8 |
| quite+a number (of)+n | 0 | 0.0 | 3 | 2.0 |
| quite+a few+n | 6 | 1.5 | 3 | 2.0 |
| article+quite+adj+n | 1 | 0.3 | 4 | 2.7 |
| quite+adj+n | 2 | 0.5 | 6 | 4.1 |

Although the most common syntactic patterns with both adjectives and nouns are identical in LOCNEC and LINDSEI-LITH, the situation regarding the patterns with *quite* and verbs (v) is very different. First of all, the pattern that is most common among NSs, *be* in any tense followed by *quite*, is used among NNS only once (see Table 20). This pattern is not followed by any other word class and is most often found at the end of an unfinished sentence or phrase as in (43), while according to previously presented table 17, learners use *quite* at the end of an unfinished sentence only seldom. Consider the following example:

(43) it was like .. it was **quite**

 [it's <X> it's yeah .. it's a city one (E02)

Table 20. Syntactic patterns with verbs

| | Number of occurrences in LOCNEC | Percentage in LOCNEC | Number of occurrences in LINDSEI-LITH | Percentage in LINDSEI-LITH |
|------------------|---------------------------------|----------------------|---------------------------------------|----------------------------|
| quite+v(Pr.S.) | 6 | 1.5 | 2 | 1.4 |
| quite+v(Past S.) | 7 | 1.8 | 5 | 3.4 |
| modal v+quite | 10 | 2.6 | 1 | 0.7 |
| be*+quite | 15 | 3.8 | 1 | 0.7 |

*be in any tense

Moreover, the most common syntactic pattern in LINDSEI-LITH is *quite+v.(Past S.)* (see Example (44)), which accounts for 3.4% of the occurrences with *quite* in this corpus. The second most frequently used syntactic pattern with verbs in LOCNEC is *modal v+quite*, but it occurs only once in LINDSEI-LITH. Thus, we can conclude that learners as well as NSs show different preferences for the syntactic patterns with *quite* and verbs.

Finally, as it can be seen from Table 21 the usage of adverbs (adv) following *quite* does not differ much in these corpora: the syntactic pattern *quite+adv* comprises 7.4% of the whole expressions containing *quite* in LOCNEC and 8.8% in LINDSEI-LITH, e.g.:

(44) (em) *I quite liked that it was small and I was able to go around with my bike (em) I could . meet my friends easily (er) (...)* (LT067)

(45) *and it was literally just sort of it built up quite quite strongly in the first sort of forty-five minutes of the film and <XX> it was just a one long series of frenetic activity <X> on and on [and on * (E27)

Table 21. Syntactic patterns with adverbs

| | Number of occurrences in LOCNEC | Percentage in LOCNEC | Number of occurrences in LINDSEI-LITH | Percentage in LINDSEI-LITH |
|-----------|---------------------------------|----------------------|---------------------------------------|----------------------------|
| quite+adv | 29 | 7.4 | 13 | 8.8 |

In addition, corpus findings by Biber et al. (1999: 545) show that among the most common collocations of adjectives with modifying adverb in conversation there are a few expressions containing *quite*: *quite good* (occurring over 50 times per million words in BrE), *quite nice* (occurring over 20 times per million words in BrE). The analysis shows that in LOCNEC as well as LINDSEI-LITH *quite good* is also most often occurring lexical collocate of *quite* and adjective with *quite nice* being the second most often used collocate used by NSs; however, learners do not use it so often: it accounts only for 1.4% of all the occurrences with *quite* in

LINDSEI-LITH (see Table 22). The analysis shows a preference for other adjectives like *sure*, *difficult*, *interesting* and *happy* among EFL learners.

Biber et al. (1999: 546) also showed that adverbs are used as modifiers of other adverbs not that often as modifiers of adjectives, which is also true for this analysis. They also found that *quite well* occurs among the most frequently used *adv+adv* collocations (occurring over ten times per million words in BrE). This collocate was found in both corpora, and it is common, especially in learners' speech. Finally, although most of the common lexical collocates in LOCNEC are found in LINDSEI-LITH as well, it could be concluded that they are more varied among NSs than among learners, as there are 14 of them found in LOCNEC and only 10 in LINDSEI-LITH. Thus, there are no occurrences of *quite a lot* (although *quite a lot of* is quite common), *sort of quite*, *quite warm* and *quite big* in learners' speech. Table 22 illustrates the findings:

Table 22. Lexical collocates in LOCNEC

| | Number of occurrences in LOCNEC | Percentage in LOCNEC | Number of occurrences in LINDSEI-LITH | Percentage in LINDSEI-LITH |
|-------------------|---------------------------------|----------------------|---------------------------------------|----------------------------|
| Quite good | 16 | 4.1 | 10 | 6.8 |
| Quite nice | 15 | 3.8 | 2 | 1.4 |
| Quite difficult | 15 | 3.8 | 4 | 2.7 |
| Quite a lot | 12 | 3.1 | 0 | 0.0 |
| Quite interested | 10 | 2.6 | 2 | 1.4 |
| Quite interesting | 10 | 2.6 | 4 | 2.7 |
| Quite a lot of | 9 | 2.3 | 7 | 4.8 |
| Sort of quite | 7 | 1.8 | 0 | 0.0 |
| Quite a bit | 6 | 1.5 | 1 | 0.7 |
| Quite happy | 6 | 1.5 | 4 | 2.7 |
| Quite well | 5 | 1.3 | 7 | 4.8 |
| Quite sure | 4 | 1.0 | 5 | 3.4 |
| Quite warm | 4 | 1.0 | 0 | 0.0 |
| Quite big | 4 | 1.0 | 0 | 0.0 |

It was also reported in Biber et al. (1999: 567) that in AmE *quite sure* occurs only in negative contexts. This applies to all sentences produced by NSs and to all, except one, produced by NNSs. In BrE it can be used in positive as well as negative sentences. The preference for the

usage common in AmE among learners is noticeable, but it is difficult to draw a conclusion here, as the frequency of *quite sure* is still quite low. Consider the following examples:

(46) er I think so **I'm not quite sure** <\B> (E03)

(47) (...) it was a very nice experience because actually when I chose the place **I I wasn't quite sure** where was where I was going I I like I googled something and (er) I saw beautiful views beautiful blue sea and (mhm) it was basically what I knew about the place (...) (LT078)

(48) (...) I know that I'm not going I'm I'm I am going to (erm) have my have English in my life . (erm) all the time and **I I I am quite sure** that my life will be outside of Li= Lithuania and most probably maybe English speaking country (...) (LT062)

3.2.2.2. Functional analysis of *quite*

As it was noted previously, the distinction between gradable and non-gradable adjectives and adverbs is important for the identification of the function of *quite*. Thus, it was found that 65.6% out of all instances with *quite* occur with a gradable adjective, adverb, noun or verb in LOCNEC, while in LINDSEI-LITH 72.8% of all occurrences of *quite* are followed by a gradable word. In addition, in LOCNEC *quite* almost always modifies a word in a positive sentence (363 out of 392 occurrences) and functions as a typical compromiser. The situation is very similar in LINDSEI-LITH (140 out of 147 are found in positive sentences). Thus, 74% of all occurrences of *quite* in LOCNEC have a lowering effect, while in 19% it seems to emphasise something, and thus could be classified as an amplifier. Similarly, 87% of all occurrences of *quite* in LINDSEI-LITH turned out to be downtoners and 11.6% are amplifiers. The semantic test presented previously was applied to all occurrences that from the context seemed to be amplifiers to prove that they in fact are, e.g:

(49) [but er I can't really afford to <begin_laughter> now <end_laughter> but er . but I'd like to again some time <\B>

 well no **I'm<?> not quite good** but I enjoy it <laughs> <\B> (E01)

(50) [you might get you might get **quite a bit of support** <\B> (E04)

First of all, in (49) it seems from the context that the word *good* should be emphasised, and thus the word *quite* has the effect of strengthening the meaning of this adjective, because what the speaker is actually saying is that he or she is not very good at something but still enjoys doing it. The ideas of being very good at something and just enjoying it are contrasted here. Moreover, the word in question is already in negation and the sentence *I'm not quite good, but I'm good to some extent* is correct, bearing in mind that *quite* can have the meaning 'to the

most absolute degree' (OED). Similarly, the utterance in (50) exemplifies a case of informal language when *a bit*, especially in British English, can mean 'a large amount' and *quite* used with *a bit* just strengthens the meaning of the word. Thus, in this case we could say that the implication is 'you might not get quite a bit of support, but you might get it to some extent'.

As the results of this part of the analysis did not reveal significant differences (both groups of speakers usually use *quite* as a downtoner), in order to test whether NNSs use *quite* for the same purposes as NSs, the functions of the sentences containing this VM were taken into consideration. In some cases it can have several functions; however, for the purpose of this study I focused only on the main function of the phrase or sentence, ignoring supplementary ones. It could be hypothesised that in most of the sentences *quite* is used when the speaker expressing some idea thinks that the word she or he is using might be too strong. Table 23 shows that it is, in fact, the function of most expressions with *quite* in LOCNEC and also one of the main functions found in LINDSEI-LITH (see Example (51)). However, in learners' corpus the function of most expressions is to describe emotions, feelings, state, his or her own experience or something personal as in (52):

(51) *but he needs to be you know he needs to convince . people to take him more seriously I I saw him in Interview with a Vampire and thought he was great in that*
 very .. erm . it's quite enjoyable but a bit long (E15)

(52) (...) *I just enjoy actually going . like <X> semester ends and you find just think about it oh I've learnt something and that's quite a nice feeling . I suppose* (LT018)

In (51) the speaker gives the reason why he or she did not enjoy the interview fully, thus perhaps when describing it, the speaker feels that the word *enjoyable* might be too strong and it is necessary to use some word scaling down the effect of the adjective. Example (52) clearly illustrates the function that is most common among NNSs, since the information provided in the sentence is related to the speaker's personal experience and the positive feeling is expressed. Table 23 shows that this function is not that common among NSs: while the occurrences with this function comprise 37.4 % of all the expressions with *quite* in LINDSEI-LITH, they comprise only 22.7% in LOCNEC. Moreover, the third most common function that is similarly used in both corpora is to describe or evaluate something as in (53), while the fourth one is to say something if the speaker does not know how to express himself or herself, or is not sure whether the particular word is really suitable or the best in that context.

(53) (...) *because the majority of films (eh) about (eh) dancing (em) are I think (em) quite superficial and this (eh) one is different (...)* (LT002)

Table 23. LOCNEC and LINDSEI-LITH: functions of the expressions containing quite

| | Number of occurrences in LOCNEC | Percentage in LOCNEC | Number of occurrences in LINDSEI-LITH | Percentage in LINDSEI-LITH |
|--|---------------------------------|----------------------|---------------------------------------|----------------------------|
| To express an idea when word is too strong | 125 | 31.9 | 52 | 35.4 |
| To describe emotions, feelings, state, his or her own experience or something personal | 89 | 22.7 | 55 | 37.4 |
| To describe or evaluate something | 84 | 21.4 | 32 | 21.8 |
| To express an idea lacking some knowledge how to do that or being not sure about the suitability of the word | 52 | 13.3 | 7 | 4.8 |
| To express an idea giving some amount of something, size, length, distance, time | 24 | 6.1 | 0 | 0.0 |
| To tell something when there is more vagueness in the sentence or the adjacent sentences | 5 | 1.3 | 0 | 0.0 |
| To compare | 6 | 1.5 | 0 | 0.0 |

Table 23 also shows that this function is much more common in LOCNEC, where the sentences with this function comprise 13.3% of all occurrences with *quite*, rather than in LINDSEI-LITH, where they comprise only 4.8%. In my opinion, the main reason for such a difference is that learners try to be more precise and avoid unfinished sentences or, in other words, do not even start a sentence if they have doubts whether they can successfully finish their thought. On the other hand, NSs are more relaxed and as their language is much more spontaneous and perhaps the tempo much faster, they freely express their ideas without bothering about some unfinished phrases or sentences as it was noted above.

Finally, there are only four functions of *quite* established in LINDSEI-LITH, while there are seven of them used by NSs. Additionally to the most commonly used functions, NSs also use *quite* in the sentences when they are aiming to specify some amount, size, length, distance or time, when there is more vagueness in the sentence or the adjacent sentences and when comparing something. Although the frequency of sentences with such functions, especially the last two, is quite low, these functions have not been found in learners' corpus at all. This might also contribute to preventing the learner from sounding more like a NS.

3.2.3. Compromiser *rather*

In this chapter the grammatical classification and characteristics of another underused compromiser *rather* (LL -13.41) by Quirk et al. (1958) and Biber et al. (1999) will be introduced. The structural analysis of *rather* will be presented in 3.2.3.1., whereas functional in 3.2.3.2.

Rather is one of the most common adverbs functioning as intensifiers that may premodify nouns and precede the determiner in this construction, especially in BrE; however, plural nouns cannot usually be premodified by *rather* (Quirk et al. 1958: 450-451). Examples (54) and (55) illustrate this usage with nouns:

(54) *He is rather a fool or rather foolish*

(55) **They are rather fools* (ibid.)

It should also be mentioned that *rather* is not used with non-gradable nouns, unless a gradable adjective occurs before it. As it can be seen from the table, this intensifier can be found in both positions: before the adjective and before the determiner. More possible constructions with *rather* mentioned by Quirk et al. (ibid. p. 446) are presented in Table 24. All the constructions, except the one in the brackets, are presented as possible with *rather* but incorrect with *fairly* or *pretty*.

Table 24. Constructions with *rather*

| Construction | Example |
|--|--|
| a comparative or too-construction | <i>rather</i> better/too small/too quickly |
| with certain noun phrases denoting adjectival qualities | <i>rather</i> a pity (<i>rather</i> pitiful is also correct) |
| with certain verbs | it <i>rather</i> annoys me that... |
| with other nouns, rather alone has variable position in relation to the indefinite article | a <i>rather</i> difficult task, <i>rather</i> a difficult task |

Quirk et al. (1958: 446) also note some other interesting differences among these downtoners. Although these three intensifiers can be used with adjectives and adverbs, *fairly* usually modifies an adjective or adverb denoting a desirable quality as in *fairly clean*, while *rather* an undesirable quality as in *rather dirty*. Another example provided by Quirk et al. for the illustration of this difference in usage is with a warm room. They claim that if while being in a warm room we feel comfortable, we say *It's fairly warm in here*, expressing the idea that 'it's warm enough'. On the other hand, if it is 'too warm', we choose the word *rather* instead of *fairly*.

Rather also occurs in the modal idiom *would rather*. This idiomatic verb is neither like the main verb, because it cannot follow other verbs, nor entirely like auxiliaries, which act as operators in negative and interrogative sentences, and the whole sequence is regarded as a separate unit. However, there are two kinds of negation in which *rather* occurs (Quirk et al. 1958: 141). Consider the following examples:

(56) *I'd rather not stay here alone.*

(57) *Wouldn't you rather live in a country?* (ibid.)

In the first kind of negation as in (56) *not* follows the whole expression, while in the second one *not* follows the first word. According to Quirk et al (ibid.), such negations are usually found in negative questions, which challenge an earlier statement or assumption. Moreover, these two kinds of negation sometimes lead to different interpretations of the whole sentence, e.g.:

(58) *Would we rather not go?*

(59) *Wouldn't we rather go?*

In (58) the speaker is asking whether it is advisable not to go, while in (59) he suggests that his or her opinion is that they would rather go and wants to hear whether the listener agrees. To add more, it is not acceptable to use the contracted verb form before *not* and it does not show active-passive synonymy (Quirk et al. 1958: 141):

(60) **I'd not rather go.*

(61) *I'd rather steal the car. *The car would rather be stolen by me.*

Interestingly, because of the ambiguity of the contraction 'd, sometimes *had rather* occurs instead of *would rather*, especially in AmE, and *should rather*, although very occasionally, is found instead of the same modal idiom (ibid. p. 142).

Perhaps the main function of *would rather* is to express preference, as it has the volitional meaning 'would prefer to', and thus it might be followed by *than*, which introduces a comparative construction:

(62) *I'd rather live in Dublin than in London.*

Moreover, one of the main *subordinators* that introduce clauses of preference is *rather than*. The function of preference is also assigned to *rather than* by Biber et al. (1999: 844). It is important to note the issue of subject-verb agreement, as a singular verb is required by grammatical concord if the first NP is singular, but in informal speech a plural verb is chosen because of the principle of notional concord, which is sometimes combined with the

proximity principle (Quirk et al. 1958: 761). Moreover, *rather than* can be interpreted as ‘and not’, because it refers to a rejected earlier assumption as in (63):

(63) *They were screaming **rather than** singing.*

Furthermore, *rather* as a conjunct according to its role class belongs to contrastives, which ‘present either contrastive words or contrastive matter in relation to what has preceded’ (ibid. p. 638), and are further subdivided into subclasses. *Rather* is found in two subclasses: reformulatory, where obviously another formulation is given, and replacive, where the speaker tends not to express his thought better but to replace what was already said by the more important information. Both types are frequently preceded by *or* as in the following example (ibid. pp. 634-639):

(64) *He was opposed by his mother **or, rather**, by BOTH his parents.*

In the given example the speaker realises that it is important that he was opposed by both parents not only his mother. Furthermore, *rather* can also appear in a collocation with *but*, emphasising the force of it. Usually, it happens if the contrast is a repudiation of what has been said or is implied by negation in the first part as in (65) (Quirk et al. 1958: 935):

(65) *I am **not** objecting to his morals, **but rather** to his manners.*

In such examples smaller constituents than clauses are linked by *but rather*.

3.2.3.1. Structural analysis of *rather*

There are 53 occurrences of *rather* in LOCNEC and only 11 in LINDSEI-LITH. Table 25 presents the frequency of the constructions discussed by grammarians as well as syntactic patterns. The differences in usage of this downtoner are clearly seen from the table. Nearly half of the occurrences (45.2%) display preference for *rather than*, which has not been found in learners’ corpus. In LOCNEC this expression is usually followed by a noun or verb. The table shows that most of the nouns either contain the definite article or no article at all. While most often *rather than+verb* include the present participle, it is often preceded by *just*. Consider the following examples (see Appendix 9 for all syntactic patterns with *rather than*):

(66) *yeah it's sort of the dip= the dip= <X> the diploma that counts **rather than the[i:] experience** . er and also the fact that French er nationals can come over to England and do a P G C E with their degree but I would<?> be hard pushed to go to France <\B>
E08*

(67) *[yeah it is with I know the card it's six pence <\B>*

 rather than t= ten pence but even so you know you're doing a lot of pages and things (...) (E11)

(68) I think it became more understandable to anybody <\B>

 because you saw the[i:] actors acting it out **rather than just reading** the play and thinking I don't understand he's using different language <\B> (E22)

Table 25. Constructions and syntactic patterns with rather, raw frequencies

| | Rather than | Would rather ('d rather) | Or rather | But rather | Rather+n | Rather+adj | Rather+v |
|--|--------------------|---------------------------------|------------------|-------------------|-----------------|-------------------|-----------------|
| LOCNEC (number of occurrences) | 24 | 8 | 1 | 0 | 7 | 6 | 3 |
| LINDSEI (number of occurrences) | 0 | 0 | 0 | 0 | 1 | 9 | 1 |

The second most often used construction among NSs is *would rather* and it is always found in a contracted form *'d rather*. In most sentences the phrase directly precedes bare infinitive (in 6 out of 9 occurrences), but there is also one occurrence in which the phrase *sort of* precedes bare infinitive, one occurrence with a pronoun preceding it and one with the adverb *like*. This construction is also not used by NNSs; however, there is one occurrence with *should rather*:

(69) (...)I think it's it is better to read a book . (er) if if you have a choice you you **should rather** read a book not to go to watch a movie in the cinema because <sniffs> once I read a book . and then I I went to watch it was Angels and Demons < laughs> (...) (LT001)

All three syntactic patterns *rather+n*, *rather+adj* and *rather+verb* are present in both corpora. Table 25 illustrates a clear preference for *rather+adj*. pattern by learners, as almost all of the occurrences appear in this pattern, e.g.:

(70) (...) I think it was . (er) . **rather impressive** (er) the one the one when I saw (er) an interpretation of Crime and Punishment here here in Lithuania in Kaunas (...) (LT006)

The other two syntactic patterns were found only in single occurrences. Thus, it seems that learners merely use *rather* before adjectives, while NSs use it in a number of different constructions and syntactic patterns.

3.2.3.2. Functional analysis of *rather*

First of all, all occurrences are found in positive sentences, thus all the phrases produced by NSs and NNSs are assertive. Table 26 proves that NNS language lacks half of the functions used by NSs.

Table 26. Functions of *rather*, raw frequencies

| | To express an undesirable quality/condition | To express preference | To express an idea with the meaning 'and not' | To reformulate what was being said |
|---|---|-----------------------|---|------------------------------------|
| LOCNEC (number of occurrences) | 15 | 19 | 15 | 2 |
| LINDSEI-LITH (number of occurrences) | 8 | 1 | 0 | 0 |

Moreover, while the most common function among NSs is to express preference, there is only one occurrence with such function in learners' speech. Here the most often used function is to express an undesirable quality/condition. In LOCNEC this function is found in the same number of occurrences as the function to express and idea with the meaning 'and not', if it is not an undesirable quality/condition and does not show any preference. The following examples express the most common functions:

(71) and people don't bother bringing them back they'd **rather** pay a fine (E11)

(function: to express preference)

(72) [yeah it is with I know the card it's six pence

 rather than t= ten pence but even so you know you're doing a lot of pages and things (...) (E11) (function: to express an idea with the meaning 'and not')

(73) (...) and this year as I mentioned I feel **rather** tired and I feel that it's like (em) the third year crisis like . it's been enough for me LT035 (function: to express an undesirable condition)

3.2.4. Compromiser *kind of*

The overview of the grammatical characteristics by Quirk et al. (1958) and Biber et al. (1999), and the discussion of vague and non-vague usage of *kind of*, the overused compromiser by NNSs (LL 4.44), will be overviewed in 3.2.4. This overview will be followed by the structural analysis of *kind of* (3.2.4.1.) and functional (3.2.4.2.). It could also be mentioned that *kinda*, a shortening for *kind of*, is even more significantly overused (LL 17.03); however, there is only

one occurrence in LOCNEC, thus *kind of* is much more appropriate for the comparative analysis.

Biber et al. (1999: 36) note that *kind of* raise serious problems for automatic analysis, because it can be used for hedging as well as part of a noun phrase. Their analysis showed that these expressions are used for both purposes in conversation. Consider the following examples (ibid.):

(74) *And I **kind of** brushed it off. (hedge)*

(75) *What sort of ideas have you come up with? I mean, do we want these **kind of** people in our team? (part of a noun phrase)*

When used for the second function, they are referred to as *species nouns* (Biber et al. 1999: 255), or as *partitive constructions*, expressing quality partition (Quirk et al. 1958: 249). Biber et al. (1999) point out that the usage of these expressions is different in conversation compared to academic prose. Thus, while superficially *kind of* can be considered a species noun in a number of occurrences, instead of establishing the precise type of something, as in example (76), the phrase might have a cardinally opposite function – introducing greater vagueness, as in examples (77) and (78):

(76) *What **kind of** little silver thing is it?*

(77) *There's a **kind of** mystery here, wasn't there?*

(78) *It's a very difficult **kind of** situation. (slightly modified examples from Biber et al. 1999: 256)*

In examples (76) and (77) *kind* can be replaced by the word *type* and still be grammatically correct; however, there is a sense of imprecision in example (77) and the speaker's intention is not to refer to the type of something but perhaps to vaguely mention something that is not concrete and somewhat unclear. Similarly, in (78) it seems that the speaker is having some difficulties and is not sure how to describe them. As it can be seen from the examples, in some cases it is very difficult if not impossible to identify the real function of the phrase and both interpretations may seem to be acceptable.

The following table presents the situation, regarding these two types of *kind of* in both corpora:

Table 27. The types of kind of in, raw frequencies

| | Total | Species nouns | Introducing vagueness | Not clear |
|--------------|--------------|----------------------|------------------------------|------------------|
| LOCNEC | 115 | 23 | 87 | 5 |
| LINDSEI-LITH | 100 | 17 | 75 | 8 |

As it can be seen from Table 27, in total there are 115 occurrences of *kind of* in LOCNEC. In some of them the phrase can clearly be replaced by the word *type*, and thus it is clearly a species noun, establishing the precise type of something mentioned after the phrase. Such occurrences will be excluded from the analysis as they do not express any vagueness. According to Biber et al. (1999: 255), species nouns behave grammatically like ordinary countable nouns. The results show that in 40 out of 115 occurrences *kind of* grammatically behaves like an ordinary countable noun: it follows a determiner, e.g. *a, the, this, that, some, any*, or any phrase that is modifying and specifying it. After taking the context into consideration, it was found that only 23 occurrences out of those 40 can be understood as meaning *type of*, and thus are species nouns. Consider the following examples:

(79) *yeah it was it was yeah to the capital I think the[i:] atmosphere you know the[i:] .. like the street cafe kind of*

 *I think it was that **kind of** atmosphere* (E07)

(80) *because .. we we don't have to do that level of maths to to understand it but it means it's harder because you you can't I haven't got the **kind of** brain that says oh that equation proves that oh [fair enough you know* (E09)

In (79) in the first sentence the speaker describes the atmosphere, thus in the following sentence *kind of* means *type of*. In (80) the phrase *the kind of brain* is also quite specific as the explanation is present in the sentence and there is no sense of vagueness. Therefore, such kind of sentences were excluded from further analysis.

The context revealed that 12 occurrences, although grammatically behaving like ordinary countable nouns, are not species nouns, as they are vague in meaning, e.g.

(81) [*you know a complete set and very very traditional . and they just had this big .. **kind of** pond in the middle of the stage and <X> they had to keep hedging round it* (E09)

(82) *at the moment .. simply because I'm interested in language acquisition it's not as a **kind of** career or anything* (E19)

The speaker of the first example is perhaps not sure whether the body of water he or she is having in mind is, in fact, called a pond, and by adding *kind of* before *pond*, shows his or her position. Similarly, in the other example the imprecision is introduced, as the speaker does not want to call it career and with the usage of *kind of* he or she strengthens this view.

It should be also mentioned that expressions with unclear usage will be excluded from the analysis. There are 5 unclear occurrences that can either be interpreted as establishing the precise type or introducing vagueness as in examples (83) and (84), or the sentence is not clear at all:

(83) *em we don't really know what it's gonna be yeah some [kind of eh community work <\B> (E55)*

(84) *for so long <XX> watched the first time when <?> about two three years old and it was black and white <XX> bang it goes into colour and <XX> tigers <XX> and all that kind of stuff <\B> (E54)*

While in (83), *kind of* can be used to emphasise that what they were going to do was just some type of community work not all the types, the sentences can also be interpreted as containing vagueness. Thus, the work mentioned in this example might not be exactly community work but just similar to this type. To add more, it seems that in (84) *all that kind of stuff* is used to define everything that is related to the phrase mentioned before the phrase *kind of*; however, the word *stuff* already introduces some vagueness, and thus when used together they create an effect of vagueness.

Similarly as in LOCNEC, in LINDSEI-LITH, 42 out of 100 occurrences grammatically behave as ordinary countable nouns; however, only half of them are species nouns and the other half are vague in usage, making 75 vague expressions in total. Consider the following examples:

(85) (...) *it was the very first time that I took part some kind of this kind of sport like (er) snowboarding and at first I was really afraid of this but at the same time very very enthusiast to learn (...) (LT004)*

(86) (...) *they were also really helpful and whenever I (er) got into a kind of problem they were always helping me and (er) (er) helped me to get out of (eh) difficult situations (...) (LT034)*

Example (85) is interesting, as it contains both types of usage: the first *kind of* shows the speaker's uncertainty, while the second one could be replaced by *type*, and thus is a species noun. Moreover, in (86) the speaker is not classifying problems but may be not sure whether

the problems he or she is having in mind were, in fact, serious enough to be called problems. Finally, the usage of 8 occurrences in LINDSEI-LITH is either not clear or seem to be interpretable in both ways as in the following example:

(87) (em) alternative music . like not not the hard and and (eh) heavy stuff like like metal
but (eh) (mm) yeah alternati= alternative rock indie yeah that that **kind of** stuff
(LT050)

The expression in example (87) is very similar as in (84), discussed above: although it seems that the phrase *that kind of stuff* is used to define the type of music, it also contains a very vague word *stuff*.

3.2.4.1. Structural analysis of *kind of*

All 23 species nouns are followed by a noun, while in LOCNEC only a small part of VMs are followed by this word class: 25 out of 87 occurrences (see Table 28).

Table 28. Word class following vague *kind of* in LOCNEC, LINDSEI-LITH, raw fr.

| | total | nouns | verbs | adjectives | adverbs | prepositions | other |
|--------------|-------|-------|-------|------------|---------|--------------|-------|
| LOCNEC | 87 | 25 | 32 | 9 | 2 | 2 | 17 |
| LINDSEI-LITH | 75 | 32 | 11 | 23 | 3 | 1 | 5 |

The most frequently used word class after *kind of* is a verb: even 32 occurrences in LOCNEC, including expressions from examples (88) and (89):

(88) [university at all and my parents were all right about that they just **kind of said**
you know well what would you like to do .. erm so it's not a problem I've ever had but I
have seen it in . you know some of my friends <\B> (E01)

(89) and . that **kind of opened** my eyes to: to .. to to . professional acting and and what
happened was really about <\B> (E04)

The second most common word class is a noun as in example (90), leaving an adjective in the third place, as there are nine sentences in LOCNEC containing *kind of* + *adj* pattern. Only two occurrences were found in *kind of* + *adv* pattern as well as two in *kind of* + *preposition*. Finally, nine occurrences were found at the end of the sentence or of some unfinished sentence or phrase, seven introducing a clause and one word following *kind of* was not found in the dictionary at all, thus it is not clear what the speaker actually meant. Consider the following examples:

- (90) *I think overall I prefer to live in town . although each has its advantages I mean when you're on campus .. everything's near everything's easy cos it's **kind of a little town** in itself isn't it [so (E07)*
- (91) *yeah but it's **kind of optimistic** at the[i:] end because . because she had a she had a young boy and . he er he loves his books .. but he loves <XXX> books and he follows him off into the distance and he says something you know (E21)*
- (92) *and erm .. we just we were up in . the <X> right at the top so we were looking **kind of straight** at the surtitles <X> have to look down to see the[i:] action (E09)*
- (93) *and their entrances<?> **kind of from** behind the[i:] audience as well so that
 [you had people<?> coming from all round you and it was really good
(E09)*

The situation with the species nouns in LINDSEI-LITH is the same: all of them are followed by nouns. Moreover, in learners' language a noun is most often used word class in sentences where *kind of* possesses a vague meaning. It could be due the fact that learners are more familiar with the usage of *kind of* as a species noun, and thus even when using *kind of* in a vague context, they still tend to follow the same pattern.

Furthermore, the usage of verbs after *kind of* is two times lower if compared to the usage of adjectives while the usage of adjectives by NSSs is quite low. A possible interpretation for this could be that learners of foreign languages are always encouraged to learn a lot of synonyms for adjectives. Thus, they know that there are many options to describe the same characteristic. It could perhaps be said that adjectives are perceived as very abstract words and vague expressions containing adjectives just strengthen that in learners' speech. The following examples illustrate the most often usage of *kind of* in terms of a word class that follows it:

- (94) *there seems to be everything seems to be more organized and more peaceful and more tidier and in here it's all **kind of a mess** at this point I don't know maybe it'll change in a few years but I don't think I'll get to experience that (LT077)*
- (95) *I think during the first year I was **kind of scared** because . I was afraid that teachers here are very strict and the rules are very strict well the rules are strict (...) (LT057)*
- (96) *(erm) I don't know when I came here it was (er) so so different from the environment in which I have been before because the people are different and (er) the lecturers are **kind of friendly** with you but in a way they you have to (erm) make the biggest effort yourself if you want to stay here (...) (LT039)*

(97) (...)once you get out the (er) this to the actual town rubbish everywhere no hygiene (erm) the food is suspicious <X> and . **kind of was** very frightening when they said you should not (er) go to any like travelling agencies that are not (er) that are not related in any way to your travelling agency because it's not safe (...) (LT018)

Adverbs and a preposition were also found after *kind of*; however, the usage of them is low as well as it is in LOCNEC: only three adverbs and one preposition. In two sentences *kind of* introduces a clause and in other three it precedes the phrase *I don't know*, e.g.:

(98) (erm) . well I thought that the city was a bit (eh) I don't know too big and too crowded I guess because . (erm) well all of the people say that you know it's the city of love and so beautiful but **I don't know it's kind of** . (eh) **I don't know** it's just not homey I guess (LT070)

In (98) *kind of* is not only followed by *I don't know*, but also precedes this phrase. *I don't know* can serve as a VM; however, it can also just simply signify absence of some knowledge. Thus, it might be that learners resort to VL when they lack some knowledge.

Having a closer look at each word class, it should be, first of all, mentioned that Quirk et al. (1958: 451) presented four different syntactic patterns that occur with *kind of* and *sort of* in informal style when the noun precedes or follows the phrase:

Table 29. Syntactic patterns with kind of and sort of with nouns (Quirk et al. 1958: 451)

| | |
|--------------|--|
| This must be | <i>a kind of</i> joke. (1) |
| | <i>kind of</i> a joke. <informal> (2) |
| | <i>a kind of</i> a joke. <more informal> (3) |
| | a joke, <i>kind of</i> . <most informal> (4) |

For this analysis only phrases used with vague meaning were taken into consideration. While the expressions that occur in construction (3) are regular: the indefinite article before and after the VM, the patterns of other occurrences are not exactly the same as the ones presented in constructions (1) and (2). Thus, in order to classify the patterns in which *kind of* is used before or after a noun, the patterns presented by Quirk et al. were slightly modified. Firstly, the expressions are considered to be following the first pattern if there is a determiner or any other modifying word before *kind of*, e.g. *a*, *an*, *this*, *some* etc. The patterns that are treated as modifications of the pattern (1) are presented in Table 30. Secondly, if there is no article or any other modifying word preceding *kind of*, the expression is classified as following pattern

(2). Finally, if a noun precedes the VM, it belongs to the group of expressions following construction (4).

Table 30. The raw frequency of the four patterns in LOCNEC and LINDSEI-LITH

| | | | | |
|--------------|-----|-----|-----|-----|
| LOCNEC | (1) | (2) | (3) | (4) |
| | 10 | 11 | 4 | 1 |
| LINDSEI-LITH | (1) | (2) | (3) | (4) |
| | 15 | 14 | 3 | 0 |

The results presented in Table 30 show that NSs as well as learners most often use the first two syntactic patterns. Pattern (3) is used quite rarely comparing to the first ones, while pattern 4 is not present at all in LINDSEI-LITH and there is only one occurrence in LOCNEC:

(99) yeah it was it was yeah to the capital I think the[i:] atmosphere you know the[i:] ..
like the street cafe **kind of** <\B> (E07)

Table 31 presents the most often found slightly modified syntactic patterns of the first pattern and displays the differences in usage of them in both corpora (see appendix 10 for all the patterns).

Table 31. Syntactic patterns in LOCNEC and LINDSEI-LITH (construction 1), raw frequencies

| | | | | |
|--------------|----------------------------|-------------------------------|-------------------------------|------------------------------------|
| LOCNEC | a <i>kind of</i> + noun | some <i>kind of</i> + noun | this <i>kind of</i> + noun | noun + <i>kind of</i> + noun |
| | 2 | 2 | 2 | 2 |
| LINDSEI-LITH | a <i>kind of</i> + noun | some <i>kind of</i> + noun | this <i>kind of</i> + noun | NP + <i>kind of</i> + noun |
| | 4 | 7 | 2 | 1 |

The indefinite article before *kind of* is used more often among learners (26.67% of the instances found in this pattern) than NSs (20%). Another noteworthy difference is the usage of *some* in front of the VM: even 46.67% of occurrences in this pattern contain this

determiner in LINDSEI-LITH, while the ones from LOCNEC comprise only 20%. Thus, learners tend to emphasise vagueness by adding the word *some*.

Table 32 shows a similar tendency regarding the definite article in construction (2): while there are two occurrences with this type of article following *kind of* in LOCNEC, there are no sentences with this pattern produced by learners.

Table 32. Syntactic patterns in LOCNEC and LINDSEI-LITH (construction 2), raw frequencies

| LOCNEC | <i>kind of</i> + a/an + noun | <i>kind of</i> + the + noun | <i>kind of</i> + noun | <i>kind of</i> + pronoun |
|--------------|---------------------------------|--------------------------------|-----------------------|-----------------------------|
| | 5 | 2 | 4 | 0 |
| LINDSEI-LITH | <i>kind of</i> + a/an + noun | <i>kind of</i> + the + noun | <i>kind of</i> + noun | <i>kind of</i> + pronoun |
| | 3 | 0 | 9 | 2 |

Moreover, as one of the first rules EFL learners get to know is a necessity to use articles with singular nouns, it could be hypothesised that in the sentences where there is no article before the VM, the article will be present before a noun if it is in singular. However, the results display that it is true only in 21.43% of the phrases, because 64.29% of the occurrences in this pattern are used without articles before a singular noun in LINDSEI-LITH, while 63.64% are used with articles by NSs. Thus, it seems that learners tend to ignore some basic grammar rules while using VL. To strengthen this point, example (100) should be taken into consideration:

(100) . (er) well . all four pictures share . **kind of** same idea but at the same time the idea is different like the setting is the same (er) I see two people (eh) sitting in a room (...)
(LT061)

Quirk et al. refers to the usage of the definite article before the words like *same* as the ‘logical’ use of *the*, because the meaning of such adjectives is ‘inalienably associated with uniqueness’ (1958: 270). In example (100) the speaker follows this rule twice; however, the student ignores it when the adjective is used after the VM.

As regards constructions with verbs, *kind of* belongs to the group of the downtoners that can precede a negative verb phrase (see example (101)). They are restricted to medial position of adverbial (*M*) in a positive clause, but can be at initial-medial position of adverbial (*iM*) in a negative one (ibid. p. 601).

(101) She **kind of** wasn’t listening (Quirk et al. 1958: 601).

The results show that almost in all occurrences *kind of* is used before a verb in a positive sentence. There are only two sentences in LOCNEC where it occurs before a verb in a negative construction (*kind of* + *didn't/doesn't*). Moreover, in most of the sentences in LOCNEC and in all sentences in LINDSEI-LITH the VM is found before a verb in present or past simple tense (see Table 33). As the frequency of the verbs in these tenses is almost equal in both corpora, we could conclude that both NSs and learners use the phrase *kind of* almost equally to talk about past and present events. In LOCNEC there are also three occurrences with modal auxiliary verbs before *kind of* and bare infinitive after it. Consider the following examples:

(102) I think it's just **kind of** . **doesn't tell** you who or when or where what <?> <\B>
(E54)

(103) (...) the movie was so so sub-far something like that (em) but anyways a little bit different because with Harry Potter there's a lot of effects an= and <overlap /> they **kind of make** for the lack of the[i:] story line <laughs> (...) (LT010)

(104) (...) I <starts laughing> guess it's kind <stops laughing> of a typical situation . . yeah I guess he **kind of embellished** on on some of her features . that's why she likes it more . guess that's about it (LT059)

(105) I think she acted very well and . again you really felt for her and erm .. you **could kind of imagine** yourself in her shoes [and <\B> (E35)

Table 33. Syntactic patterns of *kind of* in LOCNEC and LINDSEI-LITH, raw fr.

| | <i>kind of</i> + a verb in Past S. | <i>kind of</i> + a verb in Pr. S | modal aux. + <i>kind of</i> + bare infinitive |
|--------------|------------------------------------|----------------------------------|---|
| LOCNEC | 13 | 14 | 3 |
| LINDSEI-LITH | 5 | 6 | 0 |

Although the frequencies of *kind of* in other constructions rather than with verbs in simple tenses are low, they are still present in NS, but absent in NNS language (see Appendix 11 for all the patterns). Thus, learners perhaps choose only simple tenses, as they most often hear NSs use them.

3.2.4.2. Functional analysis of *kind of*

According to Quirk et al. (1958), *kind of* is used as a downtoner for adjectives, adverbs, and can apply a scale to gradable verbs, because the effect it has on the force of the verb or predication is lowering. Table 34 shows that NSs most often use *kind of* with non-gradable verbs, while learners with gradable adjectives:

Table 34. The usage of gradable and non-gradable adjectives and verbs, raw fr.

| | With gradable adjectives | With non-gradable adjectives | With gradable verbs | With non-gradable verbs |
|--------------|--------------------------|------------------------------|---------------------|-------------------------|
| LOCNEC | 11 | 7 | 4 | 27 |
| LINDSEI-LITH | 31 | 5 | 3 | 8 |

In all sentences the appropriateness of a word or phrase, following *kind of* is considered. For example, in (106) the appropriateness of *faded out* is considered and the speaker is perhaps thinking himself or herself whether this is the right word to use in such situation:

(106) and .. that **kind of** faded out and so I was in this kind of floating period <\B>
(E04)

Interestingly, if an adjective is following *kind of* and if it occurs within a NP, then the appropriateness of the whole NP is considered. In some cases the appropriateness of a longer phrase or even a clause, if *kind of* introduces it, is taken into consideration and lowering effect is applied to it. After analysing the context of the sentences with *kind of*, some functions of the phrases containing this VM were identified. The main functions are presented in Table 35 (see Appendix 12 for all the functions of the sentences with *kind of*).

Table 35. LOCNEC and LINDSEI-LITH: functions of the expressions containing kind of, raw fr.

| Functions | LOCNEC | LINDSEI-LITH |
|--|--------|--------------|
| To describe something, to call or name something, when the speaker is not sure how it should be called | 22 | 13 |
| To say something when the meaning is not completely literal | 9 | 1 |
| To tell something when there is more vagueness in the sentence or the adjacent sentences | 9 | 14 |
| To describe emotions, feelings, state, his or her own experience | 8 | 23 |
| To compare | 5 | 5 |
| To say something when the speaker doesn't know some important things | 3 | 6 |
| To express an idea when the word seems to be too strong | 3 | 16 |

As it can be seen from the table, most often NSs use *kind of* in a sentence describing, calling or naming something, when the speaker does not know how it is called. This function is also quite common among learners, but it appears only in the fourth place according to the frequency in LINDSEI-LITH. To add more, NSs also use this VM when the meaning of

something is not literal. In contrast, there is only one occurrence with such function in LINDSEI-LITH. However, in both corpora there are often a number of other VMs in the same or adjacent sentence. The most common function for learners is to describe emotions, feelings, states, his or her experience. Thus, when learners get quite personal and maybe slightly uncomfortable with the topic, they resort to VL. This function is present in LOCNEC as well; however it is not that common. The second most often used function in LINDSEI-LITH is to say something when the word seems to sound too strong for the speaker. Consider the following examples:

(107) *and what you had to do in the[i:] exam was **kind of** erm .. pick one question*

 [*sort of . compare the lighting between two of the plays you've seen* (E09)

(108) [*inside the characters so .. and and tho= those type of characters real .. or person personalities .. once you've **kind of** . got inside the character's head*

 you can develop the personality and like the[i:] erm .. the persona happy . the comedy-tragedy faces .. you can kind of go behind a mask (E04)

(109) *it was well really angering they offered us some extra V I P services which meant that . the next day they brought us some **kind of** (eh) a a slightly rotten fruits*
(LT018)

(110) *well right when I started I I didn't really like it that much I was **kind of** confused and (eh) a bit stressed and now I've I dunno I've . grown to enjoy it I guess there are (erm) there are some things that I don't like but . I dunno I think it's quite good*
(LT048)

(111) (mhm) *oh I see that it's related to painting and the painter is drawing a woman who has a **kind of** (em) . disability and when he paints her like in real she doesn't like the picture and then . (...)* (LT035)

In example (107) the speaker is trying to describe the procedures of the exam and is not sure how to do that as 'erm ..' shows. The meaning of the expression used with *kind of* in (108) is not literal, as you are not actually getting inside the character's head but just starting to understand his actions and way of thinking. Thus, this example illustrates another very common usage of *kind of* among NSs. In (109) we have the words *some* and *slightly* that are also not exact and, thus it is an example of sentences where there are more vague expressions apart the compromiser *kind of*. (110) is just an example of very common usage by learners: speaking about their emotions. Finally, in (111) the speaker avoids sounding too strong. Maybe he or she is not sure whether it is a disability or something else. To add more, in this sentence the speaker might also feel not comfortable with the topic.

4. Conclusions

In order to find out whether the learners' underuse, overuse or misuse of VL significantly accounts for the lack of naturalness and authenticity in their speech, quantitative and qualitative methods of analysis were applied. First of all, the quantitative analysis showed that the most often used VM in NS speech is *sort of*, which is significantly underused in LINDSEI-LITH in relation to LOCNEC. In contrast, Lithuanian learners most often resort to *maybe* which is only the tenth most frequent VM in LOCNEC. In addition, there are more underused VMs rather than overused in the Lithuanian corpus: 19 out of 52 are underused and 12 overused in comparison to NSs.

The qualitative analysis of underused and overused vague coordination tags and compromisers with the highest frequency in learners' corpus was twofold: structural and functional. The former revealed that, firstly, the underused tags are usually following a noun in both corpora; however, a clear preference to countable nouns among NSSs was noticed, while NNSs seem to be using half of the tags more with uncountable, and the other half with countable nouns. In addition, NSSs use tags after verbs and adjectives much more often than NNSs. Furthermore, in contrast to NNS use, NSSs in a number of instances add a tag to the whole clause instead of just a word or phrase. This use is not very common among learners. Secondly, while the learners use the only overused tag, *and so on*, mostly with nouns, the two and only occurrences of this tag in LOCNEC are after verbs.

The functional analysis of tags showed that the most common function among NSs is to indicate the imprecision of the preceding element(s). This function is also very common among NNSs; however, the preference is not that clear, because half of the occurrences of one tag were used with this function and the other half with the function of a list completer. In addition, over 80% of all the occurrences of the overused *and so on*, are functioning as list completers. Moreover, both occurrences of this tag used by NSs (there were only two occurrences in this corpus in total) are functioning as performance fillers. Some underused tags are also used as performance fillers in LOCNEC; however, instances with this function have not been found in LINDSEI-LITH. Thus, *and so on*, which is most often used in written discourse, usually does not convey any meaning when used by NSs, but NNSs perhaps use it in a similar way, as they would use it when writing.

The structural analysis showed that one of the two underused compromisers, *quite*, in more than half of the occurrences in both corpora precede an adjective, leaving nouns in the second, adverbs in the third and verbs in the fourth place. Moreover, there are fewer occurrences with

verbs and adverbs in LINDSEI-LITH than in LOCNEC. In addition, there are significantly fewer occurrences from NNS corpus which are used at the end of a sentence or unfinished phrase or introduce a clause. Furthermore, the dominant compositional patterns as regards adjectives and nouns in both corpora are *be(Pr. S.)+quite+adj*, *quite+article+adj+n*, while the two corpora show different preferences to the compositional patterns with verbs: the most common among NSs is *be(in any tense)+quite*, whereas among NNSs *quite+v(Past S.)*.

The other underused compromiser, *rather*, often precedes nouns and adjectives in LOCNEC but in LINDSEI-LITH it mainly occurs before adjectives. Moreover, the construction *rather than* is very common among NSs; however, it has not been found in LINDSEI-LITH. Furthermore, the analysis of the overused *kind of* shows that NSs use this VM most often before verbs with *kind of+v(Past S.)* and *kind of+v(Pr. S.)* as the dominant compositional patterns, while NNSs before nouns, usually in *(some) kind of+n* pattern.

Functional analysis showed that NSs usually use *quite* to express an idea when the following word is too strong, while NNSs use it to describe emotions, feelings, state, personal experiences. However, this function is also very common among NNSs. Moreover, the number of other functions with *quite* is much higher: there are only four functions found in LINDSEI-LITH and seven in LOCNEC. Similarly, *rather* is used with four different functions by NSs, while NNSs use only half of them. The dominant function in LOCNEC is to express preference, while in LINDSEI-LITH it is to express an undesirable quality/condition. In addition, NSs usually use *kind of* to describe something, to call or name something, when the speaker is not sure how it should be called, there are also other functions, but the frequencies of the occurrences with those functions are quite low. Interestingly, with *kind of*, which is significantly overused by NNSs, all functions that are present in LOCNEC are also found in LINDSEI-LITH. However, the dominant function in LINDSEI-LITH is different from LOCNEC: it is to describe emotions, feelings, states or experiences. Furthermore, the frequencies of occurrences with three other functions are quite high, when compared to the frequencies in LOCNEC. Hence, it seems that learners are using this VM in a way that is not very common among NSs.

Thus, this study largely confirms findings from earlier research (De Cock 2004, Drave 2002, Orfanò 2013), as it showed that EFL learners seem to follow the patterns common among NSs, and thus ignore or even do not know about the other patterns or functions of VMs. On the other hand, Lithuanian learners have certain specific features. For example, De Cock (2004) found that French learners underuse VMs; this study shows that although more VMs

are underused, there are certain VMs that are significantly overused. Moreover, Lithuanian learners demonstrate some untypical uses of VMs which are seldom used by NSs. A combination of those features is undoubtedly contributing to the unnaturalness of conversation. This study could only be seen as a first step to analyse spoken English produced by Lithuanian learners. Clearly, in order to get a better picture of how competence in speech develops, more research should be carried out to investigate the other aspects of spoken language. It would be useful to contrast English produced by Lithuanian learners with learners of different first languages. Undoubtedly, findings from such studies would have a huge application in EFL teaching and assessment contexts.

5. Summary in Lithuanian

Neapibrėžtumo sąvoka vartojama daugelyje sričių, pradedant filosofija ar psichologija, baigiant kalbos ir literatūros mokslais. Įdomu pastebėti, kad šios sąvokos samprata skiriasi ne tik skirtingose disciplinose, bet netgi vienos disciplinos tyrimuose. Pavyzdžiui, kai kurie lingvistai neapibrėžtumą gretina su sąšvelniais, kiti tik pabrėžia, kad šie turi panašumų arba sutapimų, o dažnai ir klasifikuoja neapibrėžtumą kaip sąšvelnio potipį. Šiame tiriamajame darbe neapibrėžtumo raiška yra suprantama kaip kalbos elementai, modifikuojantys kitus elementus, taip suteikdami jiems neapibrėžtumo ir netikslumo.

Neapibrėžtumas yra ypač svarbus gimtakalbių sakytinėje kalboje, kadangi būtent neapibrėžtumo žymikliai suteikia sakytinei, visų pirma – šnekamajai, kalbai natūralumo, spontaniškumo. Deja, kaip rodo ankstesni tyrimai, svetimkalbių kalba dažnai skamba nenatūraliai ar primena rašytinę kalbą, ir viena priežasčių – neapibrėžtumo žymiklių stoka arba netinkama vartoseną. Tad šio tiriamojo darbo tikslas yra analizuoti gimtakalbių anglų sakytinėje (šnekamojoje) kalboje vartojamus neapibrėžtumo žymiklius ir jų vartoseną palyginti su svetimkalbių (lietuvių) duomenimis. Šiam tikslui pasirinkti du tekstynai: LINDSEI-LITH, lietuvių studentų sakinės anglų kalbos tekstynas, kuris yra tarptautinio Liuveno universiteto sakinės negimtakalbių anglų kalbos vartotojų tekstyno komponentas, ir LOCNEC, analogiškas, todėl gretinamajam tyrimui tinkantis, gimtakalbių tekstynas.

52 neapibrėžtumo raiškos žymiklių dažnumas buvo patikrintas abiejose tekstynuose naudojant Antconc programą. Statistinės tikimybės testo skaičiuoklė (Log-likelihood calculator) naudota patikrinti, kuriuos neapibrėžtumo raiškos žymiklius studentai vartoja rečiau/dažniau nei gimtakalbiai. Nustatyta, kad 19 iš 52 vartojami rečiau, o 12 dažniau. Taigi tyrimas patvirtino hipotezę, kad kiekybinės neapibrėžtumo žymiklių tendencijos dviejose tirtose kalbos atmainose skiriasi. Įdomu, kad dažniausiai gimtakalbių vartojamas neapibrėžtumo žymiklis *sort of* vos keletą kartų pavartotas negimtakalbių kalboje. Tuo tarpu negimtakalbių tekстыne dažniausias yra *maybe*, kuris pagal dažnumą yra dešimtoje vietoje gimtakalbių tekстыne.

Detalesnei kokybinei analizei pasirinkti dviejų tipų neapibrėžtumo raiškos žymikliai (NŽ), kurie išskiriasi tirtose medžiagoje savo kiekybiniais dažnumo parametrais. Struktūrinė analizė leido palyginti, kaip neapibrėžtumo žymikliai sukuria gramatinį ir leksinį junglumą gimtakalbių ir lietuvių studentų šnekamojoje anglų kalboje. Šiuo požiūriu abi kalbėtojų grupės turi tam tikrų panašumų, pavyzdžiui, dauguma NŽs vartojami po daiktavardžių, o švelninantis (angl. compromiser) *quite* abiejuose tekstynuose dažniausiai eina prieš būdvardžius. Tačiau gimtakalbiai kur kas dažniau vartoja šiuos neapibrėžtumo raiškos

žymiklius samplaikose su kitomis kalbos dalimis nei tą daro negimtakalbiai. Taip pat lietuviai studentai dažniau vartoja pasakymą *and so on*, kuris gimtakalbių tekstyne rastas tik samplaikose su veiksmažodžiais, tuo tarpu lietuviai studentai dažniausiai jį vartoja po daiktavardžių.

Kitas rečiau lietuvių vartojamas švelninantis NŽ *rather* LOCNEC tekstyne dažnai eina prieš daiktavardžiais arba būdvardžius, o LINDSEI-LITH – prieš būdvardžius. Įdomu pažymėti, kad frazė *rather than* itin dažna LOCNEC tekstyne, bet visiškai nevartojama LINDSEI-LITH tekstyne. Vienintelis statistiškai dažniau vartojamas švelninantis NŽ *kind of* gimtakalbių tekstyne dažniausiai eina prieš veiksmažodžius, o negimtakalbių – prieš daiktavardžius. Taigi skirtumų nustatyta gana daug.

Šių neapibrėžtumo žymiklių funkcijų analizė taip pat atskleidė įdomių tendencijų. Pirmiausia, tyrimas patvirtino, kad lietuvių studentai nenaudoja visų funkcijų, kurios nustatytos gimtakalbių tekstyne, o kai kurias, itin retas gimtakalbių kalbos ypatybes, lietuviai išnaudoja kur kas intensyviau. Akivaizdu, jog visi šie skirtumai prisideda prie savitos ir kiek nenatūralios lietuvių studentų raiškos. Tikėtina, jog tolesni sakinės kalbos tyrimai atskleisų ir daugiau lietuvių produkuojamos anglų kalbos ypatumų, o tokių tyrimų rezultatai prisidėtų prie anglų kaip svetimšios kalbos mokymo ir vertinimo tolesnės raidos.

References

Primary sources:

De Cock, S. (ed.) 2004. LOCNEC: *Louvain Corpus of Native English Conversation*.
<https://www.uclouvain.be/en-258636.html>

Grigaliūnienė, J. & R. Juknevičienė (eds). 2011. LINDSEI-LITH: *A pilot Lithuanian component of the Louvain International Database of Spoken English Interlanguage*.
<http://www.uclouvain.be/en-cecl-lindsei.html>

Secondary sources:

Aijmer, K. 2004. Pragmatic Markers in Spoken Interlanguage. *Nordic Journal of English Studies* 3(1): 173–190.

Anthony, L. 2012. AntConc (Version 3.4.4) [Computer Software]. Tokyo: Waseda University, accessed 01 October 2015, available from <http://www.antlab.sci.waseda.ac.jp/>.

Biber, D. 1988. *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, D., S. Johansson, G. Leech, S. Conrad & E. Finegan. 1999. *The Longman Grammar of Spoken and Written English*. Harlow: Longman.

Brown, G. & G. Yule. 1983. *Discourse Analysis*. Cambridge: Cambridge University Press.

CAED – *Cambridge English Dictionary*, accessed 10 February 2016, available from <http://dictionary.cambridge.org/>

Channell, J. 1994. *Vague Language*. Oxford: Oxford University Press.

Coffin, C., A. Hewings & K. O'Halloran. 2004. *Applying English Grammar.: Corpus and Functional Approaches*. New York: Routledge.

De Cock, S. 2004. Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures (BELL)* 2: 225-246.

Dines, E. 1980. Variation in discourse – and "stuff like that". *Language in Society*. 9/1: 13-31.

Drave, N. 2002. Vaguely speaking: a corpus approach to vague language, in P. Collins, P. Peters & A. Smith (eds). *New Frontiers of Corpus Research*. Amsterdam, New York: Rodopi. 25-39.

- Grigaliūnienė, J. & R. Juknevičienė. 2011. Formulaic language, learner speech and the spoken corpus of learner English LINDSEI-LITH, *Žmogus ir žodis*, 13 (III): 12-18.
- Grigaliūnienė, J. & R. Juknevičienė. 2013. Recurrent formulaic sequences in the speech and writing of the Lithuanian learners of English. In S. Granger, G. Gilquin & F. Meunier (eds) *Twenty Years of Learner Corpus Research: Looking back, Moving ahead. Corpora and Language in Use - Proceedings 1*. Louvain-la-Neuve: Presses universitaires de Louvain. 211-222.
- Hyland, K. 1994. Hedging in academic writing and EAP textbook, *English for Specific Purposes*, 13/3: 239-256.
- Hyland, K. 1998. Boosting, hedging and the negotiation of academic knowledge, *Text*, 18/3: 349-382.
- Hyland, K. 1998a. *Hedging in scientific research articles*. Amsterdam/Philadelphia: John Benjamins.
- Hyland, K. 1998b. Persuasion and context: The pragmatics of academic metadiscourse, *Journal of Pragmatics* 30: 437-455
- Hyland, K. & P. Tse. 2004. Metadiscourse in academic writing: reappraisal, *Applied Linguistics*, 25/2: 156-177.
- Hyland, K. 2005. *Metadiscourse: Exploring Interaction in Writing*. London/ New York: Continuum.
- Hyland, K. 2005a. Stance and engagement: a model of interaction in academic discourse, *Discourse Studies*, 7/2: 173–192.
- Jankauskaitė-Jokūbaitienė, V. 2013. *Boosting Strategies in Written Learner English*. Unpublished MA thesis. Vilnius: Vilnius University.
- Jucker, A. H., S. W. Smith & T. Lüdge. 2003. Interactive aspects of vagueness in conversation, *Journal of Pragmatics*, 35: 1737-1769.
- Lakoff, G. 1972. Hedges: a study in meaning criteria and the logic of fuzzy concepts. In *Papers from the Eighth Regional Meeting, Chicago Linguistic Society. Chicago: Chicago Linguistic Society*. Pp. 183-228. (Reprinted in *Journal of Philosophical Logic* 2: 458-508, 1973)
- Log-likelihood calculator, accessed 15 December 2015, available from <http://ucrel.lancs.ac.uk/llwizard.html>

- OED – Oxford Dictionary of English. 2003. Oxford: Oxford University Press.
- Orfanò, B. M. 2013. Analysing the use of vague language in spoken interlanguage: A corpus-based study of a group of Brazilian university students learning English as a second language. In S. Granger, G. Gilquin & F. Meunier (eds) *Twenty Years of Learner Corpus Research: Looking back, Moving ahead. Corpora and Language in Use – Proceedings 1*. Louvain-la-Neuve: Presses universitaires de Louvain. 367-376.
- Overstreet, M. 2011. Vagueness and hedging, in G. Andersen & K. Aijmer (eds) *Pragmatics of Society*. Berlin, Boston: Mouton De Gruyter. 293-317.
- Quirk, R. Greenbaum, S. Leech, C. & Svartvik, J. 1985. *A Comprehensive Grammar of the English Language*. New York: Longman.
- Ruzaitė, J. 2004. Academic Precision Reconsidered: A Corpus-based Account, *SKY Journal of Linguistics*, 17: 217-247.
- Šeškauskienė, I. 2008. Hedging in ESL: A case study of Lithuanian learners. *Kalbų studijos* 13: 71-76.
- Šinkūnienė, J. 2011. *Autoriaus Pozicijos Švelninimas Rašytiniame Moksliniame Diskurse: Gretinamasis Tyrimas*. Daktaro disertacija. Vilnius: Vilniaus universiteto leidykla.
- Ullmann, S. 1962. *Semantics. An Introduction to the Science of Meaning*. Oxford: Blackwell.
- Usonienė, A. 2004. *Modalumas Anglų ir Lietuvių Kalbose: Forma ir Reikšmė*. Vilnius: Vilniaus universiteto leidykla.
- Ward, G. & B. Birner. 1993. The semantics and pragmatics of *and everything*, *Journal of Pragmatics*, 19: 205-214.
- Williamson, T. 1994. *Vagueness*. London, New York: Routledge.
- Zhang, G. Q. 2015. *Elastic Language: How and Why We Stretch Our Words*. Cambridge: Cambridge University Press.

Appendices

Appendix 1. The list of VMs in alphabetical order, raw fr.

| Nr. | VM | Raw fr. in LINDSEI-LITH | Raw fr. in LOCNEC |
|-----|------------------------|----------------------------|----------------------|
| 1. | (a) few | 31 | 91 |
| 2. | a bit of | 2 | 52 |
| 3. | a couple of | 6 | 33 |
| 4. | a little | 74 | 56 |
| 5. | a little bit | 56 | 23 |
| 6. | a lot of | 90 | 130 |
| 7. | about | 20 | 179 |
| 8. | almost | 14 | 12 |
| 9. | and all that | 3 | 7 |
| 10. | and everything | 14 | 48 |
| 11. | and so on | 29 | 2 |
| 12. | and stuff | 8 | 51 |
| 13. | and things | 2 | 85 |
| 14. | anywhere | 5 | 25 |
| 15. | approximately | 3 | 4 |
| 16. | around | 1 | 0 |
| 17. | et cetera | 0 | 4 |
| 18. | fairly | 0 | 15 |
| 19. | hundreds (of) | 1 | 3 |
| 20. | in a way | 7 | 20 |
| 21. | kind of | 75 | 87 |
| 22. | kinda | 12 | 1 |
| 23. | loads (of) | 2 | 26 |
| 24. | lots (of) | 28 | 54 |
| 25. | many | 76 | 36 |
| 26. | maybe | 291 | 86 |
| 27. | more or less | 2 | 3 |
| 28. | nearly | 0 | 6 |
| 29. | not more than | 0 | 1 |
| 30. | occasionally | 0 | 5 |
| 31. | often | 11 | 20 |
| 32. | or anybody (like that) | 0 | 1 |
| 33. | or anything | 4 | 26 |
| 34. | or so | 5 | 5 |
| 35. | or somebody | 0 | 1 |
| 36. | or something | 52 | 78 |
| 37. | or what ... | 1 | 4 |
| 38. | or whatever | 2 | 21 |
| 39. | or wherever | 0 | 1 |
| 40. | perhaps | 13 | 35 |
| 41. | probably | 108 | 102 |

| | | | |
|-----|----------------|-----|-----|
| 42. | quite | 147 | 392 |
| 43. | rather | 11 | 53 |
| 44. | round | 0 | 1 |
| 45. | several | 9 | 3 |
| 46. | some | 260 | 218 |
| 47. | somehow | 15 | 1 |
| 48. | sometimes | 57 | 43 |
| 49. | somewhere | 34 | 40 |
| 50. | sort of | 8 | 523 |
| 51. | stuff | 26 | 86 |
| 52. | thousands (of) | 2 | 1 |

Appendix 2. Raw (and normalised per 100 000w.) fr of all 52 VMs ordered according to LL value (from the most underused to the most overused VMs by learners)

| | LINDSEI-LITH | LOCNEC | LL |
|------------------------|--------------|-----------|---------|
| sort of | 8 (9) | 523 (358) | -435.17 |
| about | 20 (22) | 179 (123) | -80.79 |
| and things | 2 (2) | 85 (58) | -66.52 |
| a bit of | 2 (2) | 52 (36) | -36.73 |
| quite | 147 (163) | 392 (268) | -28.39 |
| and stuff | 8 (9) | 51 (35) | -17.61 |
| loads (of) | 2 (2) | 26 (18) | -14.44 |
| fairly | 0 (0) | 15 (10) | -14.42 |
| rather | 11 (12) | 53 (36) | -13.41 |
| stuff | 26 (29) | 86 (59) | -11.39 |
| or whatever | 2 (2) | 21 (14) | -10.45 |
| a couple of | 6 (7) | 33 (23) | -9.8 |
| or anything | 4 (4) | 26 (18) | -9.14 |
| (a) few | 31 (34) | 91 (62) | -8.91 |
| and everything | 14 (16) | 48 (33) | -6.88 |
| anywhere | 5 (6) | 25 (17) | -6.63 |
| nearly | 0 (0) | 6 (4) | -5.77 |
| occasionally | 0 (0) | 5 (3) | -4.81 |
| et cetera | 0 (0) | 4 (3) | -3.85 |
| perhaps | 13 (14) | 35 (24) | -2.62 |
| in a way | 7 (8) | 20 (14) | -1.81 |
| round | 0 (0) | 1 (1) | -0.96 |
| not more than | 0 (0) | 1 (1) | -0.96 |
| or somebody | 0 (0) | 1 (1) | -0.96 |
| or anybody (like that) | 0 (0) | 1 (1) | -0.96 |
| or wherever | 0 (0) | 1 (1) | -0.96 |
| or what ... | 1 (1) | 4 (3) | -0.77 |
| lots (of) | 28 (31) | 54 (37) | -0.57 |
| hundreds (of) | 1 (1) | 3 (2) | -0.31 |

| | | | |
|----------------|-----------|-----------|--------|
| and all that | 3 (3) | 7 (5) | -0.29 |
| often | 11 (12) | 20 (14) | -0.1 |
| more or less | 2 (2) | 3 (2) | 0.01 |
| approximately | 3 (3) | 4 (3) | 0.06 |
| or something | 52 (58) | 78 (53) | 0.18 |
| or so | 5 (6) | 5 (3) | 0.58 |
| a lot of | 90 (100) | 130 (89) | 0.7 |
| thousands (of) | 2 (2) | 1 (1) | 1 |
| somewhere | 34 (38) | 40 (27) | 1.86 |
| around | 1(1) | 0(0) | 1.93 |
| almost | 14 (16) | 12 (8) | 2.62 |
| kind of | 75 (83) | 87 (60) | 4.44 |
| several | 9 (10) | 3 (2) | 6.73 |
| sometimes | 57 (63) | 43 (29) | 14.49 |
| probably | 108 (120) | 102 (70) | 15.18 |
| kinda | 12 (13) | 1 (1) | 17.03 |
| a little | 74 (82) | 56 (38) | 18.69 |
| somehow | 15 (17) | 1 (1) | 22.38 |
| a little bit | 56 (62) | 23 (16) | 34.7 |
| many | 76 (84) | 36 (25) | 40.37 |
| and so on | 29 (32) | 2 (1) | 42.96 |
| some | 260 (288) | 218 (149) | 51.54 |
| maybe | 291 (323) | 86 (59) | 238.44 |

Appendix 3. The LL and p values for the underused VMs by NNSs

| Vagueness marker | LL | p |
|-------------------------|-----------|----------|
| et cetera | -3.85 | < 0.05 |
| occasionally | -4.81 | < 0.05 |
| nearly | -5.77 | < 0.05 |
| anywhere | -6.63 | < 0.01 |
| and everything | -6.88 | < 0.01 |
| (a) few | -8.91 | < 0.01 |
| or anything | -9.14 | < 0.01 |
| a couple of | -9.8 | < 0.01 |
| or whatever | -10.45 | < 0.01 |
| stuff | -11.39 | < 0.001 |
| rather | -13.41 | < 0.001 |
| fairly | -14.42 | < 0.001 |
| loads (of) | -14.44 | < 0.001 |
| and stuff | -17.61 | < 0.0001 |
| quite | -28.39 | < 0.0001 |

| | | |
|------------|---------|----------|
| a bit of | -36.73 | < 0.0001 |
| and things | -66.52 | < 0.0001 |
| about | -80.79 | < 0.0001 |
| sort of | -435.17 | < 0.0001 |

*the minus sign next to the LL value shows that the VM is underused

Appendix 4. The LL and p values for the overused VMs by NNSs

| Vagueness marker | LL | p |
|------------------|---------|----------|
| kind of | +4.44 | < 0.05 |
| several | +6.73 | < 0.01 |
| sometimes | +14.49 | < 0.001 |
| probably | +15.18 | < 0.0001 |
| kinda | +17.03 | < 0.0001 |
| a little | +18.69 | < 0.0001 |
| somehow | +22.38 | < 0.0001 |
| a little bit | +34.7 | < 0.0001 |
| many | +40.37 | < 0.0001 |
| and so on | +42.96 | < 0.0001 |
| some | +51.54 | < 0.0001 |
| maybe | +238.44 | < 0.0001 |

*the plus sign next to the LL value shows that the VM is overused

Appendix 5. The LL and p values for the VMs that occur with insignificant differences regarding the frequency

| Vagueness marker | LL | p |
|------------------------|--------|--------|
| perhaps | -2.62 | > 0.05 |
| in a way | -1.81 | > 0.05 |
| round | -0.96 | > 0.05 |
| not more than | -0.96 | > 0.05 |
| or somebody | -0.96 | > 0.05 |
| or anybody (like that) | -0.96 | > 0.05 |
| or wherever | -0.96 | > 0.05 |
| or what ... | -0.77 | > 0.05 |
| lots (of) | -0.57 | > 0.05 |
| hundreds (of) | -0.31 | > 0.05 |
| and all that | -0.29 | > 0.05 |
| often | -0.1 | > 0.05 |
| more or less | + 0.01 | > 0.05 |

| | | |
|----------------|-------|--------|
| approximately | +0.06 | > 0.05 |
| or something | +0.18 | > 0.05 |
| or so | +0.58 | > 0.05 |
| a lot of | +0.7 | > 0.05 |
| thousands (of) | +1 | > 0.05 |
| somewhere | +1.86 | > 0.05 |
| around | +1.93 | > 0.05 |
| almost | +2.62 | > 0.05 |

Appendix 6. Syntactic patterns with quite and adjectives

| | Number of occurrences in LOCNEC | Percentage in LOCNEC | Number of occurrences in LINDSEI-LITH | Percentage in LINDSEI-LITH |
|---------------------------|---------------------------------|----------------------|---------------------------------------|----------------------------|
| be(Past S.)+quite+adj | 61 | 15.6 | 26 | 17.7 |
| be(Pr. S.)+quite+adj | 83 | 21.2 | 41 | 27.9 |
| be(Pr. Perfect)+quite+adj | 2 | 0.5 | 0 | 0.0 |
| be(Future)+quite+adj | 1 | 0.3 | 1 | 0.7 |
| to be+quite+adj | 1 | 0.3 | 1 | 0.7 |
| quite+as+adj | 4 | 1.0 | 0 | 0.0 |
| quite+adj | 64 | 16.3 | 15 | 10.2 |

Appendix 7. Syntactic patterns with nouns

| | Number of occurrences in LOCNEC | Percentage in LOCNEC | Number of occurrences in LINDSEI-LITH | Percentage in LINDSEI-LITH |
|----------------------------|---------------------------------|----------------------|---------------------------------------|----------------------------|
| quite+article+adj+n | 23 | 5.9 | 10 | 6.8 |
| quite+article+n | 9 | 2.3 | 1 | 0.7 |
| quite+a lot of+(article)+n | 10 | 2.6 | 7 | 4.8 |
| quite+a number (of)+n | 0 | 0.0 | 3 | 2.0 |
| quite+a few+n | 6 | 1.5 | 3 | 2.0 |
| be+quite+article+adj+n | 2 | 0.5 | 0 | 0.0 |
| quite+a few+adj+n | 1 | 0.3 | 0 | 0.0 |
| quite+a few | 0 | 0.0 | 1 | 0.7 |
| article+quite+adj+n | 1 | 0.3 | 4 | 2.7 |
| quite+adj+n | 2 | 0.5 | 6 | 4.1 |

| | | | | |
|---------------------------------|---|-----|---|-----|
| article+quite+article+adj+n | 1 | 0.3 | 0 | 0.0 |
| article+quite+article+adv+adj+n | 0 | 0.0 | 1 | 0.7 |
| article+quite+article+n | 1 | 0.3 | 0 | 0.0 |
| article+quite+adj+n | 1 | 0.3 | 0 | 0.0 |
| quite+n | 1 | 0.3 | 0 | 0.0 |
| quite+article+adj+sort of+n | 1 | 0.3 | 0 | 0.0 |

**Appendix 8. Syntactic patterns
with verbs**

| | Number of occurrences in LOCNEC | Percentage in LOCNEC | Number of occurrences in LINDSEI-LITH | Percentage in LINDSEI-LITH |
|-------------------------------|---------------------------------|----------------------|---------------------------------------|----------------------------|
| Syntactic patterns with verbs | In total: 30 | | In total: 10 | |
| quite+v(Pr. S.) | 6 | 1.5 | 2 | 1.4 |
| quite+v(Past S.) | 7 | 1.8 | 5 | 3.4 |
| modal v+quite | 10 | 2.6 | 1 | 0.7 |
| have+quite+past participle | 2 | 0.5 | 0 | 0.0 |
| be+quite+present participle | 1 | 0.3 | 0 | 0.0 |
| v(Pr.S.)+quite | 1 | 0.3 | 1 | 0.7 |
| never+did+quite+v | 1 | 0.3 | 0 | 0.0 |
| don't+quite+v | 1 | 0.3 | 0 | 0.0 |
| quite+a lot+v | 1 | 0.3 | 0 | 0.0 |
| be*+quite | 15 | 3.8 | 1 | 0.7 |

*be in any tense

Appendix 9. Syntactic patterns with rather than in LOCNEC, raw frequencies

| | | | | | |
|-----------------------------------|----------------|--------------------|------------------------|--------------------------------------|------------------------------------|
| | Rather than+n | Rather than+a/an+n | Rather than+the+n | Rather than+pronoun | In total: rather than+NP |
| LOCNEC (number of occurrences) | 3 | 1 | 3 | 2 | 7 |
| | Rather than+do | Rather than+V-ing | Rather than+just+V-ing | In total: rather than+verb | |
| LOCNEC (number of occurrences) | 2 | 3 | 3 | 8 | |

Appendix 10. Syntactic patterns in LOCNEC and LINDSEI-LITH (construction 1), raw fr.

| | | | | | | | |
|--------------|-------------------------|---------------------------|----------------------------|----------------------------|----------------------------|------------------------------|-------|
| | | | | | | | Total |
| LOCNEC | a <i>kind of</i> + noun | the <i>kind of</i> + noun | some <i>kind of</i> + noun | this <i>kind of</i> + noun | NP + <i>kind of</i> + noun | very + <i>kind of</i> + noun | |
| | 2 | 1 | 2 | 2 | 2 | 1 | 10 |
| LINDSEI-LITH | a <i>kind of</i> + noun | the <i>kind of</i> + noun | some <i>kind of</i> + noun | this <i>kind of</i> + noun | NP + <i>kind of</i> + noun | all + <i>kind of</i> + noun | |
| | 4 | 0 | 7 | 2 | 1 | 1 | 15 |

Appendix 11. Syntactic patterns of kind of in LOCNEC and LINDSEI-LITH, raw fr.

| | | | | | | |
|--------------|--|---|---|---------------------------------------|---|--------------|
| | <i>kind of</i> + a verb in past simple | <i>kind of</i> + a verb in present simple | have + <i>kind of</i> + past participle | to + <i>kind of</i> + bare infinitive | modal aux. + <i>kind of</i> + bare infinitive | TOTAL |
| LOCNEC | 13 | 14 | 1 | 1 | 3 | 32 |
| LINDSEI-LITH | 5 | 6 | 0 | 0 | 0 | 11 |

Appendix 12. LOCNEC and LINDSEI-LITH: functions of the expressions containing kind of, raw fr.

| | | |
|--|---------------|---------------------|
| Functions | LOCNEC | LINDSEI-LITH |
| To describe something, to call or name something, when the speaker is not sure how it should be called | 22 | 13 |
| To say something when the meaning is not completely literal | 9 | 1 |
| To tell something when there is more vagueness in the sentence or the adjacent sentences | 9 | 14 |
| To describe emotions, feelings, state, his or her own | 8 | 23 |

| | | |
|--|---|----|
| experience | | |
| To compare | 5 | 5 |
| To say something when the speaker doesn't know some important things | 3 | 6 |
| To express an idea when the word seems to be too strong | 3 | 16 |
| To mention something that is not very important in that sentence | 2 | 0 |
| To speak about something generally | 1 | 0 |
| To make the situation not that exact or strict | 1 | 0 |
| To show his or her opinion | 0 | 2 |
| To evaluate something | 0 | 1 |
| To give a statement with the meaning 'as if' | 0 | 1 |