The role of discourse reflexivity in a linear description of grammar and discourse: the case of IMDb message boards

by

Cameron Smart

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of English
School of English, Drama, and American and Canadian Studies
The University of Birmingham
September 2013

ABSTRACT

This study investigates the role of discourse reflexivity in the linear structure of both grammar and discourse by proposing an integrated model combining an adaptation of Sinclair and Mauranen's (2006) model of Linear Unit Grammar with two of Sinclair's models for the analysis of discourse (Sinclair 1992, 1993/2004e). It is a model which can be applied both to spoken and written as well as to monologic, dialogic and polylogic discourse. In order to demonstrate the model and how it can be employed in the investigation of the role of discourse reflexivity in the linear structure of discourse, a corpus of online message board discourse of the Internet Movie Database (IMDb) website is used.

Within the scope of grammar, it is found that discourse reflexivity is particularly salient in initial suspensive elements in the linear unit and that these elements perform a variety of functions depending on the type of element sequence in which they occur and on their position in the linear structure. In discourse, it is found that there is large number of suspensive linear units between turns, i.e. those in which participants do not comply with expectations, coinciding with points where negative evaluation or antagonism is expressed. Discourse reflexivity is particularly salient at such points and therefore can be seen to play a central role in the linear structure and character of the discourse.

Dedicated to the memory of Peter ‘Sleazy’ Christopherson, Rowland S. Howard, Zoltan Varga and Don Van Vliet, all of whom died in the time it has taken me to finish this.

ACKNOWLEDGEMENTS

In presenting this thesis, I would like to express my gratitude to my supervisors Dr. Martin Hewings, Prof. Susan Hunston and Dr. Caroline Tagg for their invaluable advice and guidance over the years. I would also like to express my thanks to Jennifer Shiu for her comments on previous drafts of this thesis and to Noel Cheung for her assistance. I would finally like to thank my long-suffering family, Ana Belén Estella Cervello and Catherine Smart Estella, for their infinite patience.

CONTENTS

CHAPTER 1
INTRODUCTION

CHAPTER 2
LITERATURE REVIEW: LINEAR APPROACHES TO THE DESCRIPTION OF GRAMMAR AND DISCOURSE

CHAPTER 3
LITERATURE REVIEW: DISCOURSE REFLEXIVITY AND METADISCOURSE

CHAPTER 4
METHODOLOGY AND PROCEDURE

LIST OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

CHAPTER 8

CHAPTER 9

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 8

LIST OF SYMBOLS USED IN THE SYSTEM OF ANALYSIS

Element relations

Symbol	Description
m	Complete message oriented element
m–	Incomplete message oriented element
+m	Message oriented element which completes incomplete element
+m–	Message oriented element which partially completes incomplete element
ms	Message oriented supplement to completed message element
ms–	Incomplete message oriented supplement to completed message element
|m	Complete suspensive message oriented element
|m–	Incomplete suspensive message oriented element
|mf	Message fragment
oi	Complete interactional organizational element
ois	Supplementary interactional organizational element
|oi	Suspensive interactional organizational element
+ot–	Text focused organizational element which signals a relation between preceding and following elements
ots	A supplementary text focused organizational element
|ot	A suspensive text focused organizational element

Linear unit relations

Symbol	Description
M	Complete linear unit of meaning
M–	Incomplete linear unit of meaning
+M	Linear unit of meaning which completes previous incomplete linear unit of meaning
+M–	Incomplete linear unit of meaning which partially completes previous incomplete linear unit of meaning
MS	Supplement to linear unit of meaning
MS–	Incomplete supplement to linear unit of meaning
|M	Suspensive linear unit of meaning
|M–	Incomplete suspensive linear unit of meaning
OIS	Interactional supplement to linear unit of meaning
|OI	Suspensive linear unit of interaction

CHAPTER 1
INTRODUCTION

1.1 General aim of the thesis

This thesis presents an integrated linear model for the description of both grammar and discourse for use in both written and spoken discourse; in monologic, dialogic and polylogic discourse, i.e. discourse with more than two participants (Marcoccia 2004); and for use in computer-mediated communication (CMC) as well as more traditional paper-based and face-to-face mediums. Having presented the model, the thesis explores the role of discourse reflexivity within the linear structure of both grammar and discourse. In order to achieve this, a small corpus of online message board discourse from the *Internet Movie Database* (*IMDb*) website (http://www.imdb.com/) was compiled (hereafter the IMDb corpus) and analyzed applying the model.[1] Based on the achievements of this study, it is hoped that the integrated linear model proposed here can be applied to further discourse types both within CMC and more traditional discourse types. It is also hoped that discourse reflexivity can now be explored not simply in terms of its discourse function but also by its role in the linear structure of grammar and discourse.

1.2 Background to the thesis

1.2.1 Definitions and terminology

A linear approach to analysis is one in which the language is 'expressed as far as possible in a linear succession of units' (Sinclair and Mauranen 2006: xix). It attempts to describe

[1] The IMDb website is an online database containing information about film and television programmes, including information about the people involved, reviews etc. It is the 53rd most visited site globally with over 42 million visitors every month (from Alexa Top 500 Global Sites, http://www.alexa.com/topsites, accessed August 9th 2013).

how '[p]eople experience language as a linear phenomenon, that is, arranged along one dimension' (Sinclair and Mauranen 2006:5): time in the case of spoken language; space in the case of written language. It is an approach which has particularly been associated with John Sinclair and his associates. It can be equally applied to the description of grammar, i.e. the structure within the sentence, in written discourse, or within the move in spoken discourse, or to the description of discourse, i.e. the structure between sentences or moves. [2]

As well as describing the system of analysis developed for this study, the term *linear* description is used here to refer to three key systems of analysis: Linear Unit Grammar (LUG) (Sinclair and Mauranen 2006); the outline model for spoken discourse (Sinclair 1992) and the analytic system for written discourse structure (Sinclair 1993/2004e). [3] What is crucial in all three of these descriptions is the fact that they are envisaged as being *dynamic* descriptions of language, i.e. that the language is described by trying to reflect the real-time experience of the language user as s/he moves through the text and the type of hypotheses and mechanisms that a user would employ in dealing with the incoming gradually unfolding text.

[2] Sinclair (1985/2004d:73) uses the term *s/m* to signify *sentence/move*, a term which covers the minimum free unit of structure in either written or spoken discourse. The term will be adopted in this thesis when referring to the studies of discourse in general in Chapters 1-3. When Linear Unit Grammar is being discussed the specific term: *linear unit of meaning* will be used. The term *linear unit* is employed as roughly the equivalent of the s/m in the model of analysis used in the analysis of this study and will be used in Chapters 4-9.

[3] Note that Sinclair himself does not use the term *linear* to describe the outline model of spoken discourse or the analytic system for written discourse structure. However they are considered to be linear descriptions in this study as they meet the prerequisites of a linear description presented in Chapter 2.

The term *discourse reflexivity* was coined by Mauranen (2001, 2007, 2010, 2012), and it will be Mauranen's conception of the term which will generally be followed in this study. It is defined in this study as:

> A property evident in certain elements whereby there exists an explicit reference to the present discourse through a reference to one of the following: the text as text; a discourse event; a discourse act; the writer as text-constructor; the reader as text-decoder; a place in the text or the time or manner of the discourse act.

An example of the type of language typically classified as being discourse reflexive can be seen in Example 1.1 from the introduction of a dissertation:

> **Example 1.1**
>
> **As explained in the preceding sections**, **I would like to suggest that** the teacher paid more attention to boys than girls which was in agreement with the prevailing findings from foreign and non-foreign classrooms, but differed from the ones in ESL classrooms **(see chapter 2)**. (Farooq 2000:58)

In this example, then, two of the underlined phrases (*As explained in the preceding sections* and *see chapter 2*) explicitly direct the reader to other parts of the text. The other, *I would like to suggest that*, at the same time states that the author is making a suggestion through writing as well as labelling the upcoming clause as being a suggestion.

The term discourse reflexivity is preferred in this study to the much more widely used term *metadiscourse* principally because the term, along with the taxonomies that have been employed in its study (e.g. Vande Kopple 1985; Hyland 1998b; Hyland 2005), differ substantially from the conception of the term discourse reflexivity and the approach employed here in its description.

1.2.2 Motivations for thesis

The main motivations for this thesis as regards the linear description of language can principally be seen as taking up the agenda set by Sinclair (1992, 1993/2004e) twenty years ago for what should be the 'priorities in discourse analysis'. These were that:

a. ***A description of language be based on a small number of simple activities***

   All three of the linear systems of analysis mentioned in section 1.2.1 have a common goal: to reduce the 'vast complexity of human communicative behaviour…to a small number of simple activities' (Sinclair 1992:83) in order to better reflect a language user's experience.

b. ***An integrated description be developed***

   Sinclair (1993/2004e:85) proposes that an 'integrated description' of language be adopted, by which he means that a description based on the principles of dynamism and linearity can be developed for both spoken and written discourse by 'adapting' the outline model for spoken discourse

> and the analytic system for written discourse for that purpose. In grammar, LUG has already been proposed as a means of achieving this through capturing 'an underlying similarity in speech and writing' (Mauranen 2009b:220).

The first of these priorities is encapsulated in the description of discourse in both the outline model for spoken discourse and the analytic system for written discourse structure. Nevertheless, since their publication neither has garnered significant attention in their application. Even those studies that have investigated the analytic system for written discourse (e.g. Álvarez de Mon y Rego 2001, 2006; Moreno 2003, 2006) have tended to be selective in the categories on which they focus. No significant development of an integrated description as proposed by Sinclair in the second of his priorities has been evident. One of the main motivations of this thesis, then, is to take up the agenda of these two points set by Sinclair twenty years ago.

However, this study proposes that the model for the description of discourse can also be integrated with LUG. In LUG, Sinclair and Mauranen (2006) present a radical new grammar where traditional terminology is abandoned in favour of a limited number of elements categories defined fundamentally by the same linear principles as the outline model for spoken discourse and the analytic system for written discourse. Hence, this study proposes a description incorporating all three of Sinclair's linear models in one integrated model for both discourse and grammar.

The second main motivation for this study is related to pursuing a different approach to the study of discourse reflexivity than has typically been employed thus far, often under the name *metadiscourse*. The approach employed in this thesis differs from that typically used in two areas: firstly, in the type of text it describes and secondly in the type of description.

Metadiscourse, in contrast to Sinclair's linear descriptions of discourse, is a concept that has attracted considerable attention over the last twenty-five years or so. However, this has been almost entirely in the context of academic discourse, and, with a few notable exceptions (e.g. Mauranen 2001, 2012), specifically monologic academic discourse. If discourse reflexivity is indeed a 'discourse universal', i.e. something used by all languages, as Mauranen (2010) suggests, then the concept clearly needs to be explored in a wider range of contexts and text-types, most obviously in discourse which is non-academic and dialogic/ polylogic.

It is also true that models of metadiscourse such as Hyland (2005) and Ädel (2006) are based on taxonomies of discourse functions with little or no reference to what role discourse reflexive language plays in the structure of discourse. Thus, another motivation of the present study is to describe discourse reflexivity, not by the type of discourse functions it performs but rather in terms of its own properties and its place in the linear structure of grammar and discourse.

These motivations can therefore be summarized in three broad goals for the thesis:

1. To present an integrated linear model of both grammar and discourse which would be applicable to both spoken and written discourse;
2. To describe discourse reflexivity in terms of its own properties and its role in the linear structure of both grammar and discourse within the model;
3. To describe the role of discourse reflexivity in non-academic dialogic and polylogic discourse.

1.2.3 Data for the model

In order to achieve the three broad goals outlined in section 1.2.2, it was therefore necessary to choose a text-type and context which was dialogic and possibly polylogic but also one which would contain longer stretches of more traditional written text in order to explore the aspects of the model based on the analytic model for written discourse. It would also have to be non-academic.

The discourse type that was chosen as fulfilling these criteria was online message boards.[4] Online message board discourse is a type of web-based 'interactive written discourse' (Ferrara et al 1991) in which two or more *posters* communicate by posting messages in response to another message or *post* (Claridge 2007) in a series of connected posts referred to as a *thread* (see Figure 1.1). It is therefore dialogic and potentially polylogic.

[4] Also termed Internet / web discussion forums or fora, discussion boards, bulletin boards, electronic discussion groups, newsgroups etc. in other studies (Arendholz 2013). For the purposes of this study these terms are considered synonymous. The term *online message boards* was chosen for this study as it is the term used by the IMDb website.

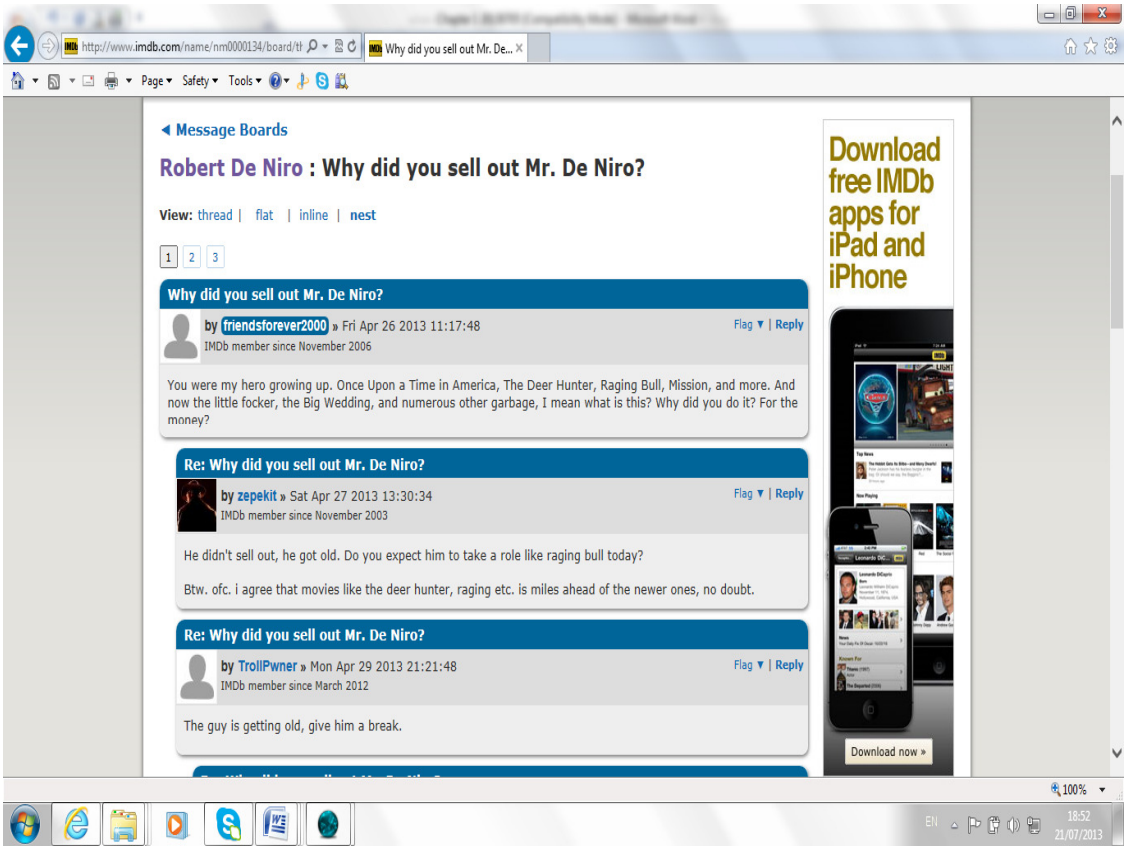


Figure 1.1: A thread in a message board in *IMDb*

One important aspect of online message boards for the purposes of this study is that, unlike chatroom communication, online message board communication is typically asynchronous, i.e. communication does not happen necessarily in 'real time' and in fact a response to a post may be posted days, weeks or months after the initial post is made. This means that posters can potentially post longer, edited pieces as they are not constrained by real-time interaction. Indeed, early characterizations of online message board discourse characterized it as possessing the characteristics of telephone

conversation combined with those of newspaper editorials and personal letters (Collot and Belmore 1996).

Message boards from the *IMDb* website were chosen as the source of the data. For the purposes of this study, the IMDb message boards provided a large amount of dialogic written discourse,[5] which is non-academic and one which potentially contains longer pieces of writing.

1.3 Objectives of the thesis

In order to reach the three broad goals for the thesis outlined in section 1.2.2, it is necessary to present the proposed integrated linear model of both grammar and discourse as an operational model. This will be achieved by its application to the discourse contained in the IMDb corpus to answer one overarching research question:

> *Do discourse reflexive (DR) elements behave in the same or in different ways in the linear structure of the IMDb corpus discourse when compared to their non-discourse reflexive counterparts?*

Given that the model provides analysis of both grammar and discourse, the question above can, in turn, can be divided into two research sub-questions:

[5] There is a large number of different message boards available on the *IMDb* website including a *Special Message Board* for every film title (over 1 million titles) and every person connected to the industry (over two and a half million) registered on the IMDb database (figures from IMDb http://www.imdb.com/ , accessed August 9th 2013).

1. *Do **DR elements** behave in the same or in different ways when compared to their non-DR counterparts in terms of the relations between elements in the IMDb corpus discourse?*
2. *Do **linear units which contain DR elements** behave in the same or in different ways when compared to linear units which do not contain DR elements in terms of the relations between linear units in the IMDb corpus discourse?*

Thus, the primary objective of this thesis is not an analysis of online message board language or CMC *per se*, but rather a presentation of the operationalization of the linear model with specific reference to discourse reflexivity. Hence, due to space limitations, a review of the literature on message board discourse or linguistic studies of CMC in general will not be provided in this thesis. It is clear, however, that there is an extensive and rapidly growing literature in this area. For example, in the area of **online message board discourse** investigations have included the study of: online message board language as a language variety using a multidimensional multi-feature model (Collot and Belmore 1996); the construction of a discourse community through message board discourse (Ho 2002); features typically assigned to CMC in online message boards (Lewin and Donner 2002); issues relating to corpus compilation of online message boards (Claridge 2007); metapragmatic utterances in online message boards (Tanskanen 2007); authorial voice in discussion groups (Hewings and Coffin 2007; Coffin et al 2012); and politeness and impoliteness in online message boards (Arendholz 2013). Additionally, research into **interaction in CMC media other than online message boards** is equally vibrant. These have included the study of: email exchanges (Harrison 1998); chatrooms (Herring 1999; Hatter 2002); mailing lists (Tanskanen 2006); blogs (Myers 2010); text

messages (Tagg 2012); online games and social media such as *Facebook* (Jones and Hafner 2012) and microblogging (Zappavigna 2012).

1.4 Outline of the thesis

Chapter 2 provides an overview of the linear approaches to the description of grammar and discourse thus far. It firstly outlines the characteristics of linear descriptions of language, namely that they are dynamic and syntagmatically-oriented and based on the syntagmatic mechanisms of prospection, completion and encapsulation. It explores the characteristics of linear descriptions of grammar, particularly LUG. It then goes on to look at the characteristics of linear approaches to the description of discourse. The most important of these are the outline for the description of spoken discourse (Sinclair 1992) and the analytic system for written structure (Sinclair 1993/2004e).

Chapter 3 looks at the studies of metadiscourse and discourse reflexivity. In particular it looks at how the two main branches of the discipline developed: the interactive model and the reflexive model of metadiscourse. It is argued that theorists face considerable challenges in trying to align the interactive model of metadiscourse to any one theoretical framework. It is also argued that the reflexive model offers in general a more theoretically sound alternative. The chapter ends by examining Mauranen's recent work on discourse reflexivity in spoken dialogic contexts.

Chapter 4 begins by describing the compilation of the IMDb corpus. It then goes on to present the model of analysis for the study in some detail. It is an analysis of two

different scopes: element relations and linear unit relations. The system of analysis of element relations is based on LUG with some significant embellishments in the shape, particularly, of suspensive elements; the system of analysis of discourse relations is based primarily on the analytic system for written structure but also incorporates features of the outline for the description of spoken discourse. In order to achieve an integrated model, fundamentally the same coding for grammar and discourse is proposed. The chapter ends by explaining how the procedures of the model of analysis manifested themselves in the *UAM CorpusTool* in the analysis of the corpus data. The means of calculating the results for the subsequent chapter is also introduced.

Chapter 5 represents the first results chapter and looks at the main three categories of the element: message-oriented, interactional-organizational and textual-organizational. The relative distribution of these is presented along with the lexical and functional properties of each category. The relative frequency of DR elements in each of these categories and their characteristics are also presented.

Chapter 6 builds on Chapter 5 by firstly looking at the element subcategories. The relative frequency of DR elements within these subcategories is also explored. It is found that DR elements are actually salient in two types of elements: initial suspensive m and oi elements. These are then explored in some detail by looking at the most salient lexical items in each and their role in the linear unit structure. A further enhancement of LUG is presented in the form of a lexico-linear means of presenting these DR elements and

element sequences. There is some indication that these DR elements contribute both to facilitate understanding and congeniality but also as tools of antagonism.

Chapter 7 is another quantitative-based results chapter. The questions posed in this chapter concern in which subcategories of linear unit DR elements are salient. The results show that they are most salient in suspensive linear units, particularly those which suspend the normal expectations between turns. It is also found that such linear units especially contain an encapsulation.

Chapter 8 is a qualitative exploration of how suspensive linear units acting between turns function in an extended extract from a single thread. It is also a means of demonstrating how the lexico-linear descriptions of elements and element sequences presented in Chapter 6 may be used in the description of an extended text. It is shown that linear units with DR elements in this position in the discourse are often chosen as tools of antagonism and play a central role in the character of the discourse.

Chapter 9 provides a summary of the achievements of the thesis. These include a multi-faceted description of the linear structure of grammar and discourse and the role of discourse reflexive elements in the IMDb corpus discourse. The achievements of the thesis are also stated in terms of the contribution to the study of discourse reflexivity, in the definition and the means provided for a structural description of the concept. It is also argued that the thesis provides significant advances in the description of linear grammar, particularly the addition of suspension and the lexico-linear description of elements and

element sequences. Finally it is argued that the thesis has provided an integrated means of describing discourse through a small number of linear unit sequence types both compliant and suspensive. The chapter ends by outlining the limitations of the study and possible further research.

CHAPTER 2
LITERATURE REVIEW: LINEAR APPROACHES TO THE DESCRIPTION OF GRAMMAR AND DISCOURSE

2.1 Overview

As was established in Chapter 1, two separate linguistic concepts are at the centre of this thesis: linearity and discourse reflexivity. The present chapter is dedicated to the review of previous research into concepts related to linearity and linear descriptions of language, particularly as presented by John Sinclair and his associates. In Chapter 3 studies related to the concept of discourse reflexivity and the associated term metadiscourse will be reviewed.

The chapter will begin with an examination of what is meant by a linear approach to language description. It will be shown that central to a linear description of language, as conceived by Sinclair, is the proposition that it is both syntagmatic and dynamic. It will be argued that central to this description are three syntagmatic mechanisms: prospection, completion and encapsulation. In section 2.2, the nature of a linear approach in general will be discussed on what has traditionally been conceived as being on two 'levels' or occupying two strata of language (Ventola 1987:58), i.e. that of grammar and of discourse, although the term different *scopes* is preferred here (Sinclair and Mauranen 2006:63) as being more congruent with a non-hierarchical linear approach. The chapter will then examine the three models which are most central to this study: Linear Unit Grammar (LUG) (Sinclair and Mauranen 2006) (section 2.3); the outline model of

spoken discourse structure (Sinclair 1992) (section 2.4); and the analytic system for written discourse structure (Sinclair 1993/2004e) (section 2.5).

2.2 Linear approaches to the description of language

2.2.1 A definition of a linear approach

As stated in Chapter 1, a linear approach to analysis is a non-hierarchical, syntagmatically-oriented description in which the language is 'expressed as far as possible in a linear succession of units' (Sinclair and Mauranen 2006: xix) in order to reflect how people experience language. According to Eggins (2004:190), Saussure (1959/1966) argues that there are two different types of relations that a sign can enter into with another: a syntagmatic relation and a paradigmatic relation. A syntagmatic relation is 'the relationship that linguistic units (e.g. words, clauses) have with other units because they may occur together in a sequence' (Richards et al 1992:369). A paradigmatic relation, on the other hand, is the relation that a linguistic unit has with another unit which could hypothetically substitute the original unit in the same place in which it occurs. The two are conventionally represented on two axes (see Figure 2.1). Thus, according to Eggins (2004:190), the syntagmatic axis 'captures the relations of sequence or chain relations between signs, and the paradigmatic axis…captures the relations of opposition or choice between signals.'

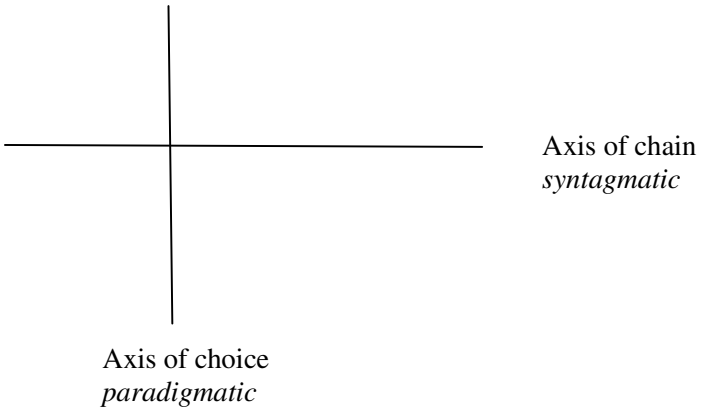


Figure 2.1: Axes of chain and choice (Eggins 2004)

A linear description, then, is syntagmatic in that it prioritizes the description of the sequence of elements, the constraints of such sequences and the mechanisms involved. This contrasts with paradigmatically-oriented descriptions such as Systemic Functional Grammar (SFG) (e.g. Halliday 1994; Eggins 2004) in which the paradigmatic choices of meaning are prioritized by being systematically contrasted. Sinclair (2003/2004g) argues that even 'conventional grammars'[6], which would seem to be syntagmatically-oriented, in that they tend to focus on how elements within the sentence can and cannot be combined sequentially, are, in fact, paradigmatically-oriented. Sinclair bases this assertion on the fact that the basic supposition of such grammars is that each word is selected on an individual basis and is chosen from a large range of paradigmatic choices of meaning.

[6] Sinclair and Mauranen (2006: xix) decline to name examples of 'conventional grammars'. Instead, they limit themselves to providing some of the general characteristics of such grammars: hierarchical; paradigmatically-oriented; based on the analysis of sentence as a complete whole; based on 'tidied up' written language etc. For the sake of having an actual example, Leech et al (1982) can be conceived as an example of 'conventional' or 'traditional' grammar in this study, complying with Sinclair and Mauranen's characterization. This is not meant to imply that Leech et al is the actual or unique grammar to which Sinclair and Mauranen refer. SFG also clearly complies with at least some of the above criteria. However, due to its rather specialized nature, SFG would not seem to fit the description of being 'conventional'. It will therefore be referred to by name in this study.

A linear description, in Sinclair's conception, is also dynamic. This dynamic perspective (Eggins 2004:51) contrasts with a 'static' (Sinclair 1985/2004d:68) or 'synoptic' (Mauranen 1996, 2009a) perspective where 'the object we are analyzing already exists in its entirety' (Brazil 1995:37). In this connection, a linear grammar has been described as a ''real-time' description of syntax' (Brazil 1995:7), in which language is seen as 'unfolding word-like element by word-like element with each element prospecting a further element until an increment is realized and a communicative need satisfied' (O'Grady 2010:86). A synoptic view of grammar, on the other hand, typically coincides with a model of grammar made up of rank scales (Halliday 1961) and constituent-within-constituent descriptions of language (e.g. that a sentence is made up of a number of clauses, which in turn is made up of a number of word groups and so on) resulting in an abstract network of hierarchical relationships by which such conventional grammatical descriptions are characterized.

In discourse, Sinclair (1985/2004d:68) argues that a dynamic model shows how the discourse proceeds 'as a continuous movement' from one point to the other. In such a description the analyst adopts the perspective of a participant experiencing the discourse as it unfolds at the point reached thus far in the discourse. This contrasts with synoptic descriptions of discourse, i.e. where the text or communicative event is considered in its entirety, e.g. in patterns of text organization (Hoey 2001), or as being made up of rank scales in a series of hierarchical relationships, e.g. Sinclair and Coulthard (1975) in spoken discourse and Mann and Thompson (1988) in written discourse. It should be noted, however, that a dynamic view of discourse is not unique to Sinclair's models of

discourse. Indeed, the dynamic/synoptic distinction is well-established in systemic-functional approaches to the description of discourse (O'Donnell 1990, 1999; O'Donnell and Sefton 1995; Eggins 2004:52). Nevertheless, the distinguishing feature of a dynamic description of discourse in Sinclair's work in comparison with such systematic-functional approaches is its focus on syntagmatic relations and the syntagmatic mechanisms evident in such relations.

2.2.2 Syntagmatic mechanisms

Characteristic of linear descriptions of both grammar and discourse is what can be termed syntagmatic mechanisms, i.e. the means by which discourse participants refer back and forward along the syntagmatic axis as they proceed through the discourse. Most prominent of these in the linear models of Sinclair and his associates are prospection, completion and encapsulation. Prospection and completion apply both to the description of discourse and grammar whereas encapsulation only applies to the description of discourse.

In grammar, such mechanisms relate to structural incompleteness/completeness. In discourse, these mechanisms are the means by which coherence is maintained (Sinclair 1993/2004e). Coherence is defined here as being the property by which the reader/listener of the text 'constructs a representation of the information it contains which integrates the propositions expressed into a larger whole' (Blakemore 2004:234) whether it be dialogic spoken discourse or monologic written discourse. Initial definitions of each syntagmatic mechanism are given here.

Prospection

Prospection 'occurs where the phrasing of a part of the text leads the addressee to expect something specific in the next part of the text' (Sinclair 1993/2004e:88) (see also Sinclair 1985/2004d, 1992, 1992/2004a, 1996/2004b; Stubbs 1996; Sinclair and Mauranen 2006). Various terms have been used to refer to similar concepts including: *prediction* (Sinclair 1981/2004c; Tadros 1985, 1994; Coulthard and Brazil 1992); *predictive assessment* (Sinclair and Coulthard 1975); *anticipation* (Winter 1977; Tadros 1985; Sinclair 1992/2004a; Brazil 1995); and *the look ahead principle* (de Beaugrande 1984). As will be seen in further discussion in the rest of this chapter, the variation in terminology often coincides with different degrees of certainty or commitment made for what is coming up later in the text.

In grammar (see section 2.3), prospection is seen in the expectations created in a text regarding the completeness or incompleteness of the grammatical structure. In spoken interactive discourse (see section 2.4), the prospection in an exchange sets the discourse framework by which a response is measured. As Sinclair (1992/2004a:12) explains:

> [Prospection] cannot determine in most cases what actually will happen, especially not in spoken interaction, but it does mean that whatever does happen has a value that is already established by the discourse at that point.

It is therefore considered to be a primary tool for controlling and managing the direction of the discourse (Sinclair 1992). In written discourse (see section 2.5), prospection in the

text can be envisaged as the means by which the author demonstrates a commitment to a future act in the text, seen, for example, in the commitment to provide *three important features* in Example 2.1.

Example 2.1

> This kind of company has three important features: the number of shareholders may be as few as two but the maximum must not exceed fifty..., a shareholder cannot transfer his shares without the consent of the company nor can any invitation be made to the general public to subscribe for shares. (Cited in Tadros 1985:18)

Completion

Completion (Berry 1981; Sinclair 1985/2004d; Sinclair and Mauranen 2006) signals the point in the text where what has been prospected in the text is fulfilled, whether it be the completion of a multi-word unit with the prospected word; the completion of a message so that it is a meaningful unit; the production of a response to an elicitation; or the production of the number of points committed to by a writer (as seen in the second part of Example 2.1) and so on. It has also been referred to as *fulfillment* or *prospected* (Tadros 1985; Sinclair 1993/2004e).

Encapsulation

Encapsulation is a retrospective labelling mechanism, which occurs in a text where an element refers to the preceding sentence/move (s/m) and in doing so encapsulates and incorporates the reference into the present s/m (Francis 1986, 1994; Sinclair 1981/2004c,

1992, 1992/2004a, 1993/2004e).[7] In written monologic discourse, this often takes the form of a noun phrase containing a demonstrative such as *this view*, as seen in Example 2.2.

Example 2.2

> At the press briefing in London during the inaugural meeting of the bank's board of governors, Henning Christophersen, vice-president of the European Commission, said: 'The EBRD must not be a political institution but plainly and simply a bank. *This view* contrasted with that of Jacques Attali, the president of the European Bank, who regards the bank's role as political and economic.'
>
> (Cited in Francis 1994:92)

As will be seen in sections 2.4 and 2.5, crucial to the concept of encapsulation is the fact that the reference is to the whole of the previous s/m as opposed to a single word or phrase.

In the next section, how having a syntagmatic orientation and a dynamic perspective manifest themselves in linear descriptions of grammar will be explored.

2.2.3 Syntagmatic relations and a dynamic perspective in grammar

A number of grammars prior to Linear Unit Grammar (LUG) can be described as being both syntagmatic and dynamic, most notably: a Grammar of Speech (Brazil 1995); Pattern Grammar (Hunston and Francis 2000); and Lexical Grammar (Sinclair

[7] See Footnote 2 in Chapter 1.

2003/2004g). The first of these represents a word-by-word linear description whereas the other two represent two different types of phraseological linear descriptions.

Unlike LUG, a Grammar of Speech (Brazil 1995) is specifically a grammar of spoken language. Its main aim is to provide a description of the dynamic process in which speakers produce language to achieve a communicative purpose through a grammatical chain. In achieving this purpose, the discourse is considered to move from an initial state, i.e. 'the special set of communicative circumstances which the speaker assumes he or she is operating in before the chain begins' to a target state, i.e. 'the modified set of circumstances that come about as a result of the listener being told what needs to be told' (Brazil 1995:48). After the first 'word-like element' of an increment has been employed, it sets up a state in which only certain ways forward are possible. The next word-like element chosen itself sets up another set of constraints as to what can follow and so on. According to Brazil, the minimum for a simple chain is a nominal (N) and a verbal (V) element:

Example 2.3

She'd been shopping

N V

When this occurs the chain is completed. However, depending on the nature of the V, an additional intermediate state or states such as the inclusion of an adverbial element or a nominal element may be required to pass through in order to reach the target state.

This simple chain, of course does not cover all possibilities in language. Brazil introduces several variations, the most important of which for this study is that of *suspension*. A suspension can be said to take place when an element occurs which is not provided for by the simple chain sequencing rules. Such suspensive elements can occur in the middle of the chain or at the beginning of the chain as seen in italics in the example provided by Brazil (1995:181) below:

Example 2.4

When they came back they asked her what was in the car.

When a suspensive element occurs it has the effect of suspending the production of an element that is permitted by the rules of the simple chain. As can be seen in the example, a suspensive element can take the form of *open selectors* (Brazil 1995:251). These comprise a variety of items including *who, when*, *because* etc., and signal that there is an obligation to eventually produce a pertinent subchain. As will be seen in Chapter 4 and beyond, the concept of suspension will play a significant role in the model of analysis proposed for this study and in the role of discourse reflexivity in IMDb corpus.

Brazil (1995) is criticized by O'Grady (2010) for such a word-by-word dynamic model, in that it ignores the 'phraseological tendency' (Sinclair 1996/2004b:31; Cheng et al 2009) in language. This tendency manifests itself in the way that words are put together and meanings are made by the combination of pre-constructed phrases and semi-pre-constructed phrases using what Sinclair (1991:100) calls the *idiom principle*, i.e. 'that a language user has available to him or her a large number of semi-pre-constructed phrases

that constitute single choices, even though they might appear to be analyzable into segments.'

One grammar which reflects this co-selection of phraseological units in a linear description is Pattern Grammar (Hunston and Francis 2000:239), in which a linear description is proposed in terms of being an overlapping flow of patterns, i.e. of individual lexical items and the syntactical environment regularly associated with them (see Figure 2.2).

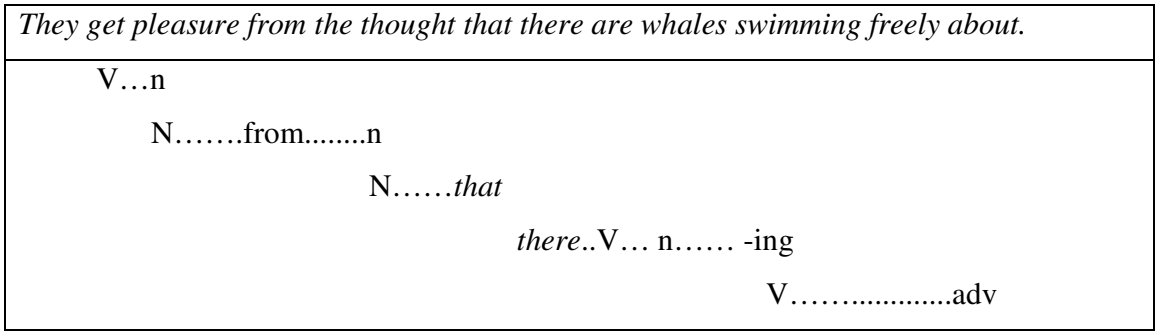


Figure 2.2: Linear description of patterns (Hunston and Francis 2000)

Unlike Pattern Grammar, which essentially employs much of the syntactical terminology of 'conventional grammars', Sinclair (2003/2004g) proposes a Lexical Grammar in which what Sinclair calls variously *extended units of meaning* (Sinclair 1996/2004b); *lexical items* (Sinclair 1998/2004f); and *meaning shift units* (Sinclair 2007 cited in Cheng et al 2009) form the basis. These lexical items are pre-constructed phrases and semi-pre-constructed phrases and demonstrate similarities with the wide range of other multi-word fixed or semi-fixed sequences from different theoretical starting points, which have been proposed over the last twenty years, e.g. *collocational frameworks* (Renouf and Sinclair

1991); *fixed expressions* (Moon 1994); *lexical bundles* (Biber et al 1999); *multi-word units* (Mason 2008); and *concgrams* (Cheng et al 2009). Sinclair argues that through such phrases, there is a great deal more co-selection of words in the use of language than has previously been acknowledged in 'conventional grammars'. This is seen for instance in the lexical item *out of the corner of my eye,* which only has two variables on the paradigmatic axis: *from* can take the place of *out of* and another possessive pronoun can replace *my* (Sinclair (2003/2004g:171). With the consequent diminished role of paradigmatic relations due to the recognition of the prevalence of such extended units of meaning, the 'special privileges of occurrence or restrictions in group structures' (Sinclair 1992/2004a:18) that are in operation for a word or phrase within its 'attendant phraseology' (Hunston and Francis 2000:2) become paramount. In a similar vein, Mason (2008) proposes a phraseological grammar in which multi-word units are derived statistically from corpus data.

How the approaches to the linear description of grammar described above relate to LUG will be discussed in section 2.3.

2.2.4 Syntagmatic relations and a dynamic perspective in discourse

Central to Sinclair's conception of a dynamic model of discourse is the proposition that there are two planes of discourse: the interactive and the autonomous planes (Sinclair 1981/2004c, 1985/2004d, 1993/2004e). The interactive plane is, according to Sinclair (1981/2004c), the plane on which the writer/speaker deals with issues relating to interaction with the outside world, i.e. 'the 'real-time' negotiation between the

participants' and therefore relates to guiding the reader/listener as to what is coming up in the text and how the part of the discourse which the participants are dealing with relates to the previous text. It is on the interactive plane that the discourse participants make 'decisions about what effect utterances should aim at, what acts they should perform or what features of the world they should incorporate' (Sinclair 1981/2004c:53). The autonomous plane, on the other hand, is where the developing record of shared experience is built up. It is where the text structure is organized and maintained. On the autonomous plane, Francis (1986:33) argues, the writer 'must provide semantic relationships between propositions or strings of propositions in such a way as to give the reader access to the shared world of accumulated meanings.' Hunston (2000:183) envisages the two planes in terms of the roles of the writer and of the reader. On the interactive plane, the writer is the text-constructor and the reader is informed of the structure of the text; on the autonomous plane, the writer is an informer and the reader is informed of the content of the text. Both planes are seen to operate simultaneously, although one plane may be more prominent than the other in any given s/m.

Such a description places at its centre the notion that discourse, whether it be spoken or written, is fundamentally an interactive process. As Sinclair (1985/2004d:68) explains:

> [t]he [written] text appears to be quite static and non-negotiable…[but]…we are assured that each reading of it, even two readings by the same reader, is a unique communicative event…a reading of a text is an event in time.

Such interaction, Sinclair points out, is more readily apparent in spoken dialogic discourse as interactional negotiations happen more explicitly due to both participants being present and active. In written discourse, or at least in written discourse as it was traditionally conceived before CMC, the participants are not both present at the same time, meaning that 'the relevant interaction is an imagined construct of the writer' (Sinclair 1981/2004c:52). As will be seen in Chapter 3, this is also a view fundamentally shared with advocates of the interactive model of metadiscourse (e.g. Hyland 2005).

The dynamism of the interaction is represented in Sinclair's description through the concept of *the text of the moment* (Sinclair 1993/2004e:82), which can be conceived as the s/m with which the reader/listener is engaged at a given time, i.e. the site of the current interaction. Thus, at the particular moment when the text of the moment falls upon a particular s/m, both planes of discourse are in operation. As the reader/listener moves onto the next text of the moment the previous text of the moment passes onto the autonomous plane where the new material is worked into existing knowledge, records are updated and stored. The next s/m then becomes the text of the moment with both planes in operation again. In this way, the discourse moves from the negotiable meaning of the potential communicative event on the interactive plane to the non-negotiable artifact of the autonomous plane. The shared knowledge on the autonomous plane is retrievable whereas the actual language used, generally, according to Sinclair, is not. It is a process which Sinclair (1993/2004e:91) calls 'complete textual erasure', i.e. one in which the wording of the previous texts of the moment is lost to the reader. As Auer (2005:75) points out, 'pragmatic experiences' can be held in the memory for a long time. In contrast

our memory 'seems to disattend form'. Hence, a previous s/m beyond the immediately prior s/m is considered to have become part of the information on the autonomous plane and the related information is 'no different from another, non-linguistic experience' (Sinclair 1992/2004a:13) in the past.

Table 2.1: Features prominent on the interactive plane

	Features	**Description**	**Example**
a.	Predictions	A commitment made by the writer at one point in the text to perform another subsequent act of discourse	*There are five types of…*
b.	Anticipating	As above but without the commitment, but which allows the author to perform a subsequent act	*Fruit drinks usually contain high quantities of sugars…*
c.	Self-reference	When the proposition concerns the text itself	*This book is…*
d.	Discourse labelling	When the acts are named as they occur	*Heat is defined as…*
e.	Participant intervention	When the author adopts directly his participant status on the interactive plane	*We allow wide margins for error.*
f.	Cross-references	Alternatives to linear sequence in a text	*See Table 6.3 above**

* No example provided by Sinclair

Initially, Sinclair (1981/2004c:54) proposes six different categories of features prominent on the interactive plane (as seen in Table 2.1). Some of these categories eventually feature, with some modification, in aspects of the outline model of spoken discourse structure (Sinclair 1992) and categories of the analytic system for written discourse structure (Sinclair 1993/2004e). It can also be noted in passing that, according to

Mauranen (1993a), certain categories, particularly those of *self-reference*, *discourse labelling* and *participant intervention* coincide with the concept of discourse reflexivity as defined in Chapter 1. This will be returned to in Chapter 3.

The terms *prediction* and *anticipation* (as seen in categories a. and b. in Table 2.1) coincide with Tadros (1985, 1994) and both fall within what can generally be termed prospection (see section 2.2.2). *Prediction* is defined by Tadros (1994:70) as 'an interactional phenomenon – a commitment made by the writer to the reader, the breaking of which will shake the credibility of the text.' *Anticipation*, on the other hand, is associated with a more general guessing of what is coming up in the text. Prediction is based on a model in which a pair is made up of two members: *the predictive member* (V); and *the predicted member* (D). Tadros proposes a taxonomy of different subtypes of prediction as seen in Table 2.2.

Table 2.2: Categories of prediction (Tadros 1985, 1994)

Category	**Explanation**	**Examples**[8]	
		V member	**D member**
Enumeration	V member carries a signal that commits writer to enumerate (e.g. numeral, discourse self-reference nouns etc.)	*This kind of company has three important features:*	*the number of shareholders may be as few as two..., a shareholder cannot transfer his shares without...*
Advance labelling	Writer labels and commits to performing a	*It is important to distinguish between*	*Nominal wages are wages in terms of*

[8] Examples taken from both Tadros (1985) and Tadros (1994)

	discourse act.	*real and nominal wages.*	*money, the term money wages is perhaps...*
Reporting	Sentence contains one report structure and one proposition attributed to others.	*It has been rightly said...*	*…that every commodity is nothing more than a bundle of services.*
Recapitulation	A member predicts an evaluation by recalling information from earlier in the text.	*We have said that the underlying representations, lexical as well as phonological are abstract...*	*There is, however, one very obvious sense in which underlying representations are more abstract...*
Hypotheticality	A member signals a detachment from the real world by creating a hypothetical world.	*Suppose the legislator could draft rules that were absolutely wrong...*	*As it is, legal uncertainty is counterbalanced by judicial flexibility…*
Question	V has interrogative syntax, committing the writer to abandon detachment at later point and answer it.	*Is college worthwhile?*	*Education is one of society's most profitable measurements.*

Each category of prediction as seen in Table 2.2 is defined in terms of a series of linguistic criteria which have to be met in order to be included in the categorization. For instance, Example 2.1 above exemplifies an enumeration.

Another model which employs Sinclair's dynamic model of discourse is Francis (1986, 1994), in which two types of labels: *advance* and *retrospective labels* (also called *cataphoric nouns* and *anaphoric nouns* respectively (Francis 1986)) are presented. These are unspecific nominal groups which require the co-text to achieve lexical realization.

These are roughly the equivalent of *signalling nouns* (Hoey 1983; Flowerdew 2003); *carrier nouns* (Ivanič 1991); *shell nouns* (Hunston and Francis 2000; Schmid 2000); *textual nouns* (Álvarez de Mon y Rego 2001); and *unspecific anaphoric nouns* (Yamasaki 2008). They also represent one type of *Vocabulary 3* items (Winter 1977).

Advance labels are seen by Francis as providing predictive and organizing functions and so can be conceived as being similar to the sub-technical and discourse self-reference nouns seen in Table 2.2, for instance, in the V member of the category of enumeration *there are three reasons:*. The second type of label, retrospective labels, represents an encapsulation, as seen in section 2.2.2. They are, according to Francis (1986), a central means by with the author expresses his/her evaluation of the previous text. Such labelling nouns have been found as crucial in the development of authorial stance (Charles 2003; Mur-Dueñas 2003–2004).

As stated in section 2.2.2, the major criterion for the identification of a retrospective label or encapsulation, according to Francis, is that it does not refer to any specific nominal group, rather, that it represents a replacement for a clause or group of clauses, guiding the reader as to how the reader should interpret the stretch of discourse by providing an appropriate label (see Example 2.2 above). This is contrasted by Sinclair (1993/2004e:84) with what he calls 'point-to-point' cohesion, which is classified as the type of cohesion that refers to less than one sentence such as a pronoun referring to a noun. Point-to-point cohesion plays a large role in, for example, Halliday and Hasan's (1976) model of cohesion. Sinclair argues that, whereas cases of encapsulation have a clear structural role

in a dynamic model, 'point-to-point' cohesion does not. Moreno (2006) provides some empirical evidence to provide some validation for this theory. Based on an experiment with a group of readers of a text it is concluded that those syntagmatic mechanisms seen in section 2.2.2 are genuinely textual whereas it was found to be unnecessary and unusual for a reader to search through the text looking for point-to-point cohesive ties in order to make sense of the text.

In written discourse, Sinclair (1993/2004e:83) describes encapsulation as the means by which the previous sentence relinquishes its role as text of the moment on the interactive plane and thus removes its discourse function. It is therefore a 'cancellation mechanism' (Crompton 2006:257) which 'cancels [the previous sentence's] interactive force' (Sinclair 1992/2004a:15) leaving simply the meaning that has been created. Hence, in the process of encapsulation the previous sentence is pushed onto the autonomous plane to be incorporated into the accumulated information and therefore loses its interactive role thus signalling the end of its interactive role and re-affirming its role as shared knowledge. Encapsulating devices can therefore be seen as provoking *plane change*, i.e. the move from being on the interactive to the autonomous plane. As Sinclair (1981/2004c:56) explains:

> By referring to a preceding utterance with discourse labels like *question* or *reply*, a speaker or writer encapsulates the old interaction in this new one, and the discourse proceeds, in a sense talking about itself. There are many signals of this operation as well as discourse self-reference items…They all share the ability to refer to the preceding

> utterance as merely a stretch of language, recognizing in some way its interactive force but not necessarily attending to it.

Again, in passing we can see that plane change as described by Sinclair here, would seem to be carried out by linguistic items that are discourse reflexive (e.g. *question*, *reply*) according to the definition provided in Chapter 1. In regards to discourse reflexivity, Francis further divides retrospective labels into metalinguistic and non-metalinguistic labels (see Table 2.3). A metalinguistic label is described as a nominal group which labels a stretch of discourse as a linguistic act, e.g. *this argument* or *that statement* (Francis 1994:83). Such metalinguistic labels would seem therefore to fit the definition of discourse reflexivity as provided in Chapter 1 although, as will be seen in Chapter 3, they do not typically feature in models of metadiscourse.

Table 2.3: Advance and retrospective labels (Francis 1994)

	Type	**Definition**	**Examples**
Metalinguistic nouns	Text nouns	Nouns which refer to the formal textual structure of discourse	*Phrase, question, sentence, word, excerpt, page, paragraph, passage, quotation, section, term*
	Language activity nouns	Nouns which refer to some language activity or result	*Account, contrast, controversy, debate, example, instance, reference, squabble, theme*
	Illocutionary nouns	Nouns which represent the nominalization of verbal processes, usually acts of	*Accusation, advice, allegation, announcement, answer, claim, complaint, comment, conclusion, excuse, statement, suggestion*

		communication	
	Mental process nouns	Nouns which refer to cognitive states and processes and the results thereof.	*Analysis, assessment, attitude, belief, idea, opinion, theory.*
Non-metalinguistic nouns		Nouns which refer to concepts, states of affairs etc. in the outside world	*Approach, area, aspect, case, matter move, problem, stuff, thing, way.*

The rest of this chapter will be dedicated to looking at three linear descriptions, firstly of grammar in LUG and secondly of discourse in the outline model of spoken discourse structure and the analytic system for written discourse structure.

2.3 Linear Unit Grammar

2.3.1 Aims and background

LUG is a dynamic, syntagmatically-oriented description of language, which, according to Sinclair and Mauranen (2006), is as applicable to spoken language as it is written language. Its most salient feature is the use of chunking in order to define the units of analysis: *the element* and *the linear unit of meaning*. Additionally, unlike A Grammar of Speech and Pattern Grammar (section 2.2.3) it is a grammar which disregards all recognizable terminology associated with conventional grammars in favour of two broad categories of elements: message-oriented (M) and organizational-oriented (O) elements.

LUG has directly influenced several studies since the initial samples of analysis were provided by Sinclair and Mauranen. These have, for the most part, been in relation to spoken language, and have included: the study of phraseology in non-native spoken

language (Cheng 2007; Cheng et al 2009); tense and aspect in a variety of World English (Van Rooy 2008); rhetorical structures in native and non-native university lectures (Mauranen 2009a); phraseology in English as a lingua franca (ELF) (Mauranen 2009b); the use of discourse particles in spoken discourse and text book representations of spoken discourse (Lam 2010); uncooperativeness in spoken discourse in ELF (Carey 2011); and formulaic organizing chunks in spoken and written ELF (Carey 2013). It has also been employed in the study of critical discourse and phraseology of both spoken and written discourse (Milizia and Spinzi 2008).

2.3.2 An overview of the system of analysis

The classification of elements used in LUG is based primarily on their role in sequential relations as briefly summarized in Table 2.4.

Table 2.4: Summary of the system of analysis of LUG

Element type	Status	Symbol	Description
Message increment	Core status elements	M	Complete message-oriented element
	Qualified status elements	M–	Incomplete message element
		+M	Completion of message element
		+M–	Partial completion of message element
		MS	Supplement to message element
		MS–	Incomplete supplement to message element
		MR	Revision to message element
		MA	Message adjustment
		MF	Message fragment

Organizational		OI	Interactional organizational element
		OT	Textual organizational element

Sinclair and Mauranen propose a five-step analysis of the text in question as outlined in Table 2.5. The first of these stages is concerned with the chunking of the text; Steps 2–4 with the classification and sub-classification of these chunks into elements; and the final stage allows further analysis to occur through the recombination of the elements to create linear units of meaning. These stages represent a general movement from the 'pre-theoretical notion' of the *chunk* (Sinclair and Mauranen 2006:51) up to the recombining of elements ready for the requirements of more abstract grammatical descriptions.

Table 2.5: Five-step analysis for LUG

Step	Description
1	The assigning of Provisional Unit Boundaries (PUBs) to the text based on chunking
2	The classification of each provisional unit as either: message-oriented (M) or organization-oriented (O)
3	The division of O elements into two sub-types: OI (interactive organizational) elements and OT (text-oriented organizational) elements
4	The division of M elements into sub-types based on how substantial or complete they are (M, M–, +M–, MS, MS–, MR, MA, MF)
5	Recombination of the elements into linear unit of meaning for further descriptive analysis

Issues relating to these different steps will be discussed in sections 2.3.3–2.3.5.

2.3.3 Chunking in Linear Unit Grammar

Step 1

The key procedure in Step 1 is the establishment of what Sinclair and Mauranen call Provisional Unit Boundaries (PUBs) through the process of chunking the text. Sinclair and Mauranen (2006:6) conceive chunking as 'a natural and unavoidable way of perceiving language text as it is encountered.' The term *chunk* itself dates back to Miller (1956) and refers to the process whereby the short-term memory processes incoming language by dividing it into small chunks of language. The length of these chunks is dictated by the capacity of the working memory for processing information with the 'magical' number for the working memory being 'seven, plus or minus two'. Sinclair and Mauranen cite studies from a variety of disciplines to back their claim both relating to the reception and production of language, such as Abney (1991:257), who argues that intuitively the following sentence is read 'chunk at a time':

Example 2.5

[I begin] [with an intuition]: [when I read] [a sentence], [I read it] [a chunk] [at a time].

This, Abney claims, is partly based on prosodic patterns, with the strongest stress of the sentence falls on one chunk and pauses occur between chunks, and partly on lexico-syntactical considerations whereby the chunk typically comprises one single content word 'surrounded by a constellation of function words'. In terms of language production, Sinclair and Mauranen cite studies in psycholinguistics which have described spoken language production as being divided into fairly short unit fragments (Levelt 1989; Mukherjee 2001). Cheng et al (2008) provide some validation for the claim that chunking is evident in both language reception and production in investigating the coincidence of

PUBs derived from the division of the written transcript of spoken discourse and the tone units of the same discourse. They conclude that there is a high correlation between tone units in spoken discourse and LUG chunks and that the same chunking process is at work in both.

Following from Abney above, Sinclair and Mauranen (2006:130) argue that chunking is a pre-theoretical process. Indeed, they argue that some variation between different language users or analysts is to be expected as different people are likely to process the incoming text in different ways. This, according to Mason (2008), reflects an underlying belief on the part of Sinclair and Mauranen that there is no uniform grammar which all users apply but rather that each individual applies his/her own resources to deal with the incoming text based on his/her own experiences and knowledge. Nonetheless, Sinclair and Mauranen (2006: 167ff.) do provide some anecdotal evidence of inter-rater agreement. Mason criticizes such an intuitive approach as remaining 'unsatisfactory' and instead proposes the use of a multi-word unit algorithm to derive the chunks through a more 'objective means'.

An *element* is the term used to describe a chunk once it has been classified employing the system of analysis described in Steps 2–4 below. It is therefore a procedural-based unit rather than the abstract linguistic units which feature in other descriptions. The boundaries of each element are determined by its unique place in the discourse and this may vary from occasion to occasion depending on the conditions in place. Nevertheless, proponents of Emergent Grammar theory argue that such 'procedural chunks' (Bybee

1998) will contain many examples of pre-existing, pre-constructed regularities borrowed from previous discourses which are retrieved from the memory to deal with the necessities of the present discourse. These chunks are generally unstable but may move towards being fixed into recognizably structural units depending on the how repeated their use is (Bybee and Hopper 2001). As Sinclair and Mauranen (2006:39) argue, through repetition, chunks become entrenched in our long-term memory and the more fixed a multi-word expression becomes the more likely it is that it will be processed as one entity and therefore more likely to occupy one chunk. In doing so, Sinclair and Mauranen are adhering to what Stubbs (1996:41) calls the seventh principle of linguistics in the neo-Firthian tradition, i.e. that 'Much language is routine.' Thus, from both an Emergent Grammar and from a Neo-Firthian point of view, it is likely that chunks in LUG will often coincide with such established multi-word units.

2.3.4 Message and organization-oriented elements

Sinclair and Mauranen present the classification of these elements as a series of binary choices that the researcher must make. These decisions are based on various factors including the orientation of the element, i.e. towards sharing information or towards organizing the discourse; the type of syntagmatic relations that the element displays with the preceding or following element and so on (see Table 2.4).

Step 2

In Step 2 elements are classified in terms of their orientation, i.e. whether the element is primarily message-oriented (M), that is oriented towards incrementing the 'shared

knowledge' of the participants as the discourse unfolds or whether it is oriented towards the organization (O) of the discourse (see Table 2.6 for examples). Message-oriented elements are seen by Sinclair and Mauranen as functioning on the autonomous plane (as described in section 2.2.4) as they are primarily related to the adding to and maintaining of the shared records of the participants.

Step 3

In Step 3, O elements are divided in terms of what is called, their *focus*, i.e. whether they are focused on facilitating the interaction or the understanding of the relations between parts of the text. These are termed OI (interactive organizational) elements and OT (text-oriented organizational) elements respectively. OI elements, are those elements whose function is to 'manage turn-taking, changes of topics and the interrelations among chunks of contents and stretches of discourse' (Sinclair and Mauranen 2006:60), whereas OT elements are those whose function it is to manage the internal coherence of the text.

Table 2.6: Summary of the system of analysis with examples [9]

Element type	**Status**	**Symbol**	**Examples**
Message increment	Core	M	*I liked that in him* M
	Qualified	M–	*in certain cases / you can* M– / +M
		+M	*he announced /two reviews* M– /+M
		+M–	*From now on, / the Liberal Democrats/ have to present themselves*

[9] All examples from Sinclair and Mauranen (2006)

			M– / +M– / +M
		MS	*Internal reviews / into the party's structure / and communications* +M / MS / OT/ MS
		MS–	*that wants power/ and /knows/ what it wants to do* MS / OT/ MS– / +M
		MR	*FROM HAMBURG TO GEDSER / I THINK / A PLACE CALLED GEDSER* MS/ OI / MR
		MA	*It was a famous Estonian tele- / television* MA / +M–
		MF	*IT /WE / WE CROSSED TO DENMARK* MF / MF / M
Organizational		OI	*I mean; I think; I'm not sure; or something; right; some kind of; erm; yeah; ha; like; yes yes; mhm; oh dear*
		OT	*and; as well as; because; but; first of all; or; for example; now; so; too*

OI elements in LUG (see Table 2.6 for examples) are therefore similar to interactional elements detailed in other frameworks, such as *discourse markers* (e.g. Schiffrin 1987; Fraser 1990, 1999; Biber et al 1999; Müller 2005); *discourse particles* (Stenström 1994; Aijmer 2002); *pragmatic markers* (Brinton 1996; Erman 2001); *discourse signals* (Stenström 1990); *discourse connectives*, *discourse operators* and *cue markers* (Blakemore 2004); *frame*, *focus* and *starter* acts (Francis and Hunston 1992); as well as a variety of grammatical features identified by Biber et al (1999) as being typical of conversation, including: *filled pauses*; *interjections*; *attention signals*; *prefaces*; *tags*; and *vocatives*. Most pertinently for this study, what Sinclair and Mauranen term OI elements

have featured in interpersonal or interactional metadiscourse (e.g. Vande Kopple 1985; Hyland 2005) in interactive models of metadiscourse. These will be discussed in section 3.3.

These language features have been described as having a variety of discourse functions in discourse depending on the theoretical framework in which they are based, including, according to Sinclair and Mauranen: initiating and maintaining interaction; control timing; structuring the interaction using framing and focusing moves; managing response strategies, such as moving towards desirable outcomes; opening and closing of turns and interactions and so on. Mauranen (2009b:223) divides OI elements into two categories: convergent, i.e. those elements 'where participants indicate effort and willingness to maintain the collaborative activity of speaking' by, for example, keeping the floor, showing listenership etc. (e.g. *yeah*, *ok*); and divergent, where elements indicate a shift in direction of the discourse usually by framing a new opening or indicating a speaker change (e.g. *in my point of view*).

OT elements, on the other hand, are seen as focusing on the text rather than the interaction. These have variously been called *linking words* (Chalker 1996); *conjunction* (Halliday and Hasan 1976); *Vocabulary 1* and *Vocabulary 2* (Winter 1977); *logical operators* (Sinclair 1993/2004e:83), as will be seen in section 2.5.2; and *text connectives* and *transition markers* in interactive models of metadiscourse (e.g. Vande Kopple 1985; Hyland 1998b; Hyland 2005), as will be seen in Chapter 3. They are conceived as being explicit textual signals of what have been variously termed: *conjunctive relations*

(Halliday and Hasan 1976); *semantic relations* (Crombie 1985); *clause relations* (Winter 1994) and *relational propositions* (Mann and Thompson 1986).

In terms of their role on the planes of discourse, OI elements clearly function on the interactive plane as facilitators of interaction. The case of OT elements is less clear, however. These have been described as having a 'dual status' (Mauranen 1993a:151), functioning both at the level of propositional content on the autonomous plane and at the level of 'textual expression' on the interactive plane. This will be discussed in relation to models of metadiscourse and discourse reflexivity in section 3.3.3.

Sinclair and Mauranen (2006:59) concede that the distinction between M and O elements 'is intuitively clear to speakers and also reflected in many linguistic models, but hard to pin down in precise terms' due, in part to the 'multi-functional nature of many elements'. A case in point is the distinction between elements which contribute to the propositional content by providing some sort of stance or comment and those which are used for more pragmatic and/or interactive uses. Whereas LUG forces the analyst into making a dichotomous decision, previous research has conceived this distinction as being on a continuum: from a literal meaning with some interactional qualities (e.g. *I guess*, *probably*); to those which are mainly interactional or pragmaticalized (*I see*, *you know*, *I mean*); to those that are purely interactional (*Aha*, *mhm* etc) (Stenström 1990; Brinton 1996).

Similarly, as Sinclair and Mauranen (2006:72) indicate, the distinction between OT and OI elements may also be difficult to maintain as many of these items have textual as well as interactional functions. This has been corroborated in studies of discourse markers such as *well* (Müller 2005; Lam 2010). Similarly, it has been argued that elements which traditionally have been associated with a textual orientation have strong interactional qualities as seen, for example, in the study of the word *so* (Müller 2005; Bolden 2009); and vice versa in the case *of course* (Thompson and Zhou 2000). Such difficulties in the exact delimitation of these types of organizational elements led Siepman (2005) to advocate the abandoning of the distinction altogether.

2.3.5 Types of message-oriented elements

Step 4

In Step 4, M elements are subdivided according to several different criteria, namely, *status*, *linear expectations*, *substantialness* and *interpenetration*. The last two of these are distinctions made in the description of spoken discourse. As this thesis focuses on written discourse, insubstantial elements (MF) and overlapping elements (MA and MR elements) will not be discussed here due to space restrictions. Examples are provided, however, of each in Table 2.6.

Table 2.7: Classification of elements in LUG[10]

	Element orientation								
	Action-oriented		**Message-oriented**						
	Focus		***Status***						
	Situation-focused	**Text-focused**	**Core status**	**Qualified Status**					
				Linear expectations					
				Prospecting			**Additional**		
				Substantial		**Insubstantial**	**Prospected**	**Unexpected**	
				Interpenetration				***Interpenetration***	
				Overlapping	**Separate**			**Overlapping**	**Separate**
OI	✓								
OT		✓							
M			✓						
MA				✓					
M–					✓				
MF						✓			
+M							✓		
MR								✓	
MS									✓
MS–					✓				✓
+M–					✓		✓		

[10] The element type +MF also appears in some of the analyses in Sinclair and Mauranen, but does not appear in the main presentation of the element types. It is therefore not included here.

M elements are firstly divided according to their *status*, i.e. whether the element contains a complete meaning element in itself or if it is incomplete or dependent on another element to provide complete meaning. This distinction is expressed in the categorization of elements as: *core M element* and *qualified M elements*. A *core M element* is one with a complete meaning in itself which does not require anything else to complete its meaning. It therefore does not prospect and is not prospected:

Example 2.6

I DIDN'T REALLY ENJOY BEING ON THE BOAT

(Cited in Sinclair and Mauranen (2006:108))

A core M element displays none of the syntagmatic mechanisms discussed in section 2.2.2. All remaining M elements are classified as qualified M elements, i.e. elements in which the meaning is not complete and which rely on other elements to provide full meaning.

Firstly, qualified M elements are divided in terms of their linear expectations, i.e. the division of elements in terms of whether they are prospective or retrospective. These are termed *prospecting* and *additional* M elements respectively. Prospecting M elements include M–, MF, MA, +M–, MS– elements. As was seen in section 2.2.2, prospection is considered to be one of the three main syntagmatic mechanisms in a dynamic model of language. An example of a prospecting M– element can be seen in Example 2.7:

Example 2.7

the train goes	M–
on the ferry	+M

(Cited in Sinclair and Mauranen (2006:152))

The final group of qualified M elements is that of additional elements. This category contains two types of M elements. The first of these are prospected +M elements (as seen in Example 2.7). These provide completion (see section 2.2.2) for the sense of semantic-structural incompleteness left by the prospection in the M– element (Sinclair and Mauranen 2006:136).

The second type of additional elements is that of unexpected additional elements (i.e. MS and MS– elements). These are elements which are not prospected and yet are not considered to be core elements. Rather, their quality is that of providing something supplementary to the already established core. The exact nature of the syntagmatic mechanism in place in MS elements is not specified other than that such elements 'do not start something new' (Sinclair and Mauranen 2006:84). However, they can be divided into two subtypes. The first of these provide a supplementary specification of the core:

Example 2.8

WE CROSSED TO DENMARK	M
IN THE MORNING	MS

(Cited in Sinclair and Mauranen (2006:108))

The second type of MS elements are what Brazil (1995:121) calls *reduplications*, i.e. a further element in an extension of the linear unit of meaning which is of the same kind as

the previous element, where the analyst is required to recognize syntactic and/or semantic parallels between elements. An example of this type of MS element is seen in Example 2.9 where *waiters* and *beggars* are MS elements providing further examples.

Example 2.9

I like that in him	M
polite to old women	MS
like that	MS
and	OT
waiters	MS
and	OT
beggars	MS
too	OT

(Cited in Sinclair and Mauranen (2006:121))

2.3.6 The Linear Unit of Meaning

Step 5

The final stage proposed by Sinclair and Mauranen is a series of operations carried out in order to produce a Linear Unit of Meaning (LUM). A LUM is described as a *topic increment* (Sinclair and Mauranen 2006:150) consisting of a core M element and any dependent M elements (i.e. +M, +M–, M–, MS, MS–, MR and MA), as such, they are the rough equivalent of Brazil's *increment*. These operations entail the removal of *non-standard elements,* i.e. OI, MA, MF, and MR elements at the same time as making a note of their position in the discourse (Sinclair and Mauranen 2006:96).

Thus, Sinclair and Mauranen argue that they have proposed a procedure whereby any text can be processed from the pre-theoretical process of chunking through to its recombination 'with the minimum possible disturbance to the linear sequence' (Sinclair and Mauranen 2006:6). The process has been criticized by Mason (2007:2) for being little more than 'a pre-processing step for a conventional syntactic analysis' which fails to produce much more than 'a mere filter that turns messy spoken dialogue into 'grammatically well-formed' sentences'. Sinclair and Mauranen do make a suggestion as to how LUG may be pushed towards the more phraseological approaches to linear grammar described in section 2.2.3, by categorizing M elements as either *textual objects* (i.e. a word representing one of the main traditional word classes (nouns, adjectives etc.) either on its own or as the head of a phrase e.g. *the ferry*) or textual incidents (a combination of two or more textual objects, e.g. *I wondered*; *from one island*). However, this remains tentative.

It could be argued, then, that questions remain regarding the validity and exact nature of LUG at each step, whether it be the exact theoretical underpinning of the concept of chunking; the lack of evidence of inter-rater agreement; the reliance on the analysts' judgment and intuition; the exact delimitation of the categories of elements; and the ultimate use of the whole process in terms of providing insights into the description of language, however intuitively sound the coding process seems. Not all of these can be addressed in this thesis. However, it is hoped that this thesis will demonstrate that rather than abandoning the categorization of elements seen in LUG (M, OI, OT etc.) for more

conventional hierarchical grammatical terms, the coding can be usefully maintained and extended in the description of grammar and also be extended to the description of discourse.

2.4 The outline model of spoken discourse structure

2.4.1 Aims and background

The outline model is a model proposed by Sinclair for the analysis of general conversation. It is what Sinclair calls a 're-prioritization' of the better known Sinclair-Coulthard model (Sinclair and Coulthard 1975), which was specifically developed to analyze dialogic classroom spoken discourse. In the Sinclair-Coulthard model 'syntagmatic patterns of discourse' (Sinclair and Coulthard 1975:29) are expressed in terms of a system of analysis based on Speech Act theory (Austin 1962; Searle 1969); with an initiating *elicitation* move prospecting a *response* and so on. In subsequent adaptations, particularly Tsui (1994), this led to an increasingly paradigmatic orientation providing systems in which the paradigmatic choices available to speakers at certain junctures are governed by differing *pragmatic presuppositions* (Tsui 1994:162), i.e. those background beliefs the speaker takes for granted to be true when s/he initiates an exchange. Sinclair (1992) argues that this had led to an overemphasis on the situation, i.e. on the 'non-linguistic conditions' (Brazil 1995:173), in which the utterance occurs. The outline model, in contrast, represents a model 'where the higher patterns (of language) can be described without reference to any particular social use' (Sinclair 1992:88). Thus, in the outline model, the text itself provides everything necessary for its own

interpretation with the syntagmatic mechanisms of prospection and encapsulation as the guiding mechanisms. It is summarized by Sinclair (1992:87) as seen in Figure 2.3.

Basic structures

I-prospects-R	=	I
Challenge	=	C
I-without-R	=	I*

I R (F)
I* (F)
I C = I…
I* C = I…
I R C = I…

Where I= Initiation; R = Response; C= Challenge; F= Follow up

Figure 2.3: Outline model for spoken discourse (Sinclair 1992)

This will be discussed in sections 2.4.2–2.4.4 in relation to the syntagmatic mechanisms introduced in section 2.2.2.

2.4.2 Prospection

In spoken dialogic discourse, according to Sinclair, prospection is employed as a means of managing the future direction of the discourse; something deemed necessary due to possible differing social intentions of participants. Such management is possible as '[e]ach initiation prospects that the utterance following it will be interpreted under the same set of presuppositions as the initiation itself' (Sinclair 1992:83), thus placing the next utterance within the 'discourse framework' (Burton 1980) created by the prospection in the previous utterance. As can be seen in Figure 2.3, an I move prospects an R move whereas an I* move does not. In terms of the original Sinclair-Coulthard model, a

prospecting I move would be seen most typically in an eliciting move. In Example 2.10, the second speaker's utterance will be interpreted in principle as a response to the elicitation provided by the first speaker, thus, allowing the initiator to achieve his/her purpose through the completion of the exchange (Brazil 1995:41).

Example 2.10

Why? Did you wake up late today?	I
Yeah, pretty late.	R
Oh dear.	F

(Cited in Francis and Hunston (1992))

2.4.3 Encapsulation

In the outline model, encapsulation is seen as occurring in the Follow-up (F) move in the exchange structure and in Challenges (see section 2.4.4). Sinclair argues that there are situations when the F move may be obligatory, particularly if it is required to make sure that all parties are confident that the expected communication has taken place or in a didactic context (as the original Sinclair-Coulthard model was), where an evaluative follow-up is necessary. In other contexts it is optional.

Up to this point, the F move had been defined in terms of what it was not, in that, to use the outline model's terminology, it was defined in terms of neither containing a prospection nor of being prospected (Coulthard and Brazil 1992:71) or as being a non-predicting, non-predicted and non-initial move (Stubbs 1983). In the outline model, in contrast, the F move is defined by the linguistic mechanism it contains, i.e. an

encapsulation of the previous move or moves in the exchange. The reference present in the encapsulation may be explicit, for example in phrases containing a deictic reference e.g. *That's a shame* (see section 2.5.2) or less explicit, such as a minimal verbless item like *Hah*, *Mm* or *Oh dear* (see Example 2.10). These, according to Sinclair are sufficient to indicate that the speaker's understanding of the exchange up to this point and that the discourse is proceeding coherently. As can be seen in Example 2.10, the encapsulation can also provide an evaluation of the rest of the exchange.

2.4.4 Challenges

One of the major criticisms of the original Sinclair-Coulthard model is that it is inadequate in dealing with everyday conversations as it is based on a co-operative model of discourse (Burton 1980, 1981; Eggins and Slade 1997). The supposition in such a model is that all participants are working together towards the joint production of the discourse. Yet, as Burton (1980, 1981) points out, spoken discourse outside the classroom is characterized by numerous moments when participants insult each other, argue, try to assert themselves, refuse to do what they have been told to do and so on. Burton (1981:70) therefore proposes two possible responses to an initiating move:

1. A supporting move by politely agreeing, complying or supporting the discourse and presuppositions of the previous move; or
2. A challenging move by not agreeing, not supporting or complying with the presuppositions and additionally by counter-proposing, ignoring or telling the original speaker that his/her opening was misguided, badly designed and so on.

Supporting moves, therefore, facilitate the discourse by complying with the syntagmatic constraints, whereas challenging moves hold up progress by not complying. Such challenging moves are incorporated into several adaptations of the Sinclair-Coulthard model, such as Burton (1980, 1981), Tsui (1994), and Warren (2006); in Francis and Hunston (1992) as *protest* acts; and in O'Donnell (1990) as *suspensions*.

A Challenge (C) is treated in the outline model as a special type of initiating move. By definition, a challenge does not fulfill the presuppositions of the previous utterance and so initiates a new exchange. For instance, if the second speaker in Example 2.10 had chosen not to respond to the elicitation, as had been prospected and had responded as in the invented Example 2.11, this would be judged in terms of not fulfilling the prospection and would therefore be classified as a challenge.

Example 2.11[11]

Why? Did you wake up late today?	I
Depends what you mean by late.	I (challenge)

As Sinclair argues, a challenge contains an encapsulation and thus cancels the interactive function of the previous utterance or pair of utterances leaving the utterances' contributions on the autonomous plane. It is again worth noting, in the context of this study, that the challenge in Example 2.11 contains the discourse reflexive verb *mean*. As Sinclair (1992:87) comments, in such challenges 'the subject matter becomes the

[11] Invented example

discourse itself'. The syntagmatic mechanisms, their role and function on the interactive plane are summarized in Table 2.8.

Table 2.8: Summary of the role of syntagmatic mechanisms in the outline model

Element of exchange structure	**Syntagmatic mechanism in operation**	**Role of syntagmatic mechanism**	**Function on interactive plane**
Initiation (I)	Prospection	Places next utterance in discourse framework so that it will be interpreted in terms of the same set of presuppositions as initiation.	Overtly contributing to the management of discourse on interactive plane through prospection.
Response (R)	Completion	Fulfills prospection made in I and opens up the possibility of subsequent encapsulation.	By fitting with presuppositions of I, R operates on interactive plane to some extent.
Follow-up (F)	Encapsulation	An explicit or non-explicit reference which encapsulates the previous I and R thus providing support for the exchange by: checking that participants have agreed on function of I and R; commenting on exchange as it stands; reacting to R in the context of I; providing evaluative feedback	Overtly contributes to the management of discourse on interactive plane through encapsulation, thus cancelling the discourse function of I and R and consigning them to the autonomous plane and terminating the exchange.
Challenge (C)	Possible Prospection	Being an I move, it may contain a prospection as described for an I above.	As Initiation above
	Encapsulation	An explicit or non-explicit reference which encapsulates the previous I or R thus challenging the presuppositions	As Follow-up above

Initiation (I*)	~	~	~

2.5 The analytic system for written discourse structure

2.5.1 Aims and background

The analytic system for written discourse structure (Sinclair 1993/2004e) represents an exploratory application of the dynamic model of discourse (Sinclair 1981/2004c) described in section 2.2.4. Thus Sinclair sets out to demonstrate that a small number of relationships between sentences in operation on the interactive plane may account for the coherence in a text.

Hence, in the analytic system, the text of the moment, is described in terms of how it relates to the text up to the previous sentence and what it prospects in the following sentence. In doing so, it borrows from other syntagmatically-orientated descriptions, particularly prediction in text (Tadros 1985, 1994) and advance and retrospective labels (Francis 1986, 1994) (see section 2.2.4). The relations between the text of the moment and the surrounding sentences are principally described in terms of one default mechanism: *encapsulation*; one variation of the default mechanism: *prospection*; and two principal exceptions to the default mechanisms: *verbal echo* and *overlay*. These will be examined in turn.

2.5.2 Encapsulation

Sinclair (1993/2004e:83) argues that:

> [t]he default hypothesis [for written text organization] is...that there is an underlying structure to discourse where each new sentence makes reference to the previous one, and encapsulates the previous sentence in an act of reference.

As can be seen in Table 2.9, encapsulation is divided by Sinclair into two subtypes, *deictic* and *logic*. Deictic encapsulation here is essentially what Francis calls retrospective labels (see section 2.2.4), i.e. what has previously been termed *discourse deixis* or *text deixis* (Levinson 1983:85) as it refers to a portion of the same discourse. Sinclair (1993/2004e) provides examples of *neutral pro-forms* (Ribera 2007), e.g. *this* and *that* and determiner + noun phrases, such *as*: *This very obvious ethos* and *this subject*. However, Sinclair also includes less explicit references within the category of deictic encapsulation.

Example 2.12

(2.2)This is the first generation in history to delude itself into thinking that because one particular language, English, seems to be very widely understood, no other language need to be learned.

(3.1) The foreign language requirement in the UK's National Curriculum will help change **things** a bit.

(Cited in Sinclair (1993/2004e:98)

According to Sinclair, *things* in Example 2.12 is a 'lexically weak word' which carries out a deictic encapsulation of the previous sentence despite the absence of a demonstrative or a determiner in the noun phrase. As we have seen in section 2.4, in

spoken discourse, Sinclair conceives *Mm* and *Hah* as well as the verb phrase *you mean* as being specific enough to encapsulate the previous exchange and this would seem to be consistent with this wider view of deixis and encapsulation in general.

The other type of encapsulation presented in the analytic system for written discourse structure is logical encapsulation, which manifests itself most frequently in what Sinclair calls, *logical operators,* such as *and*, *however*, *by contrast*, *rather* etc., i.e. OT elements in LUG (see section 2.3.4). Whereas such elements are conventionally seen in terms of providing some sort of link between parts of the text, Sinclair (1993/2004e:83) describes them as 'progressively determining the status of a previous sentence in relation to the current one'. Hence, *however* is not conceived by Sinclair as simply linking two sentences, but rather as meaning 'notwithstanding a previously stated position (PSP) in the previous sentence'; *by contrast* is conceived as meaning 'by contrast with the PSP' and so on. In essence, then, it can be argued that logical encapsulation is conceived by Sinclair in terms of being a type of elliptic deictic encapsulation where the deictic reference to *the PSP in the previous sentence* is inferred by the reader.

Table 2.9: Summary of the analytic system for written discourse structure

Type of act of reference	Explanation	Subtypes	Explanation	Examples
Encapsulation	When the 'text of the moment' (TOTM) refers to the previous sentence as the new sentence takes over the status of 'state of text' and the previous sentence relinquishes the same	Logical	When the TOTM signals the status of the previous TOTM in relation to the present e.g. ellipsis, logical operators and inferred	*...with all of them, British people will wish – and need – to communicate.* ***And yet****, how often do we see British commercial representatives abroad hard put to communicate with their potential customers...*
		Deictic	When the TOTM labels the previous TOTM and includes it as an element of the present TOTM	*A Finnish manufacturer would not dream of using Finnish to market a product in Germany or France...* ***This very obvious ethos*** *is not going to change with 1992*
Prospection	When the phrasing of TOTM leads the addressee to expect something specific in the next sentence	Topic selection	When the addressee can predict that the next sentence will pick up the new topic and develop it	***The single market will make trading conditions even more competitive****.* *As the rewards for enterprise increase, so businesses will have to refine...*
		Attribution	When quoted speech is introduced	*It included* ***the statement*** *'En el mercado unico todo el mundo habla various idiomas',...*
		Advance labelling	When the addressee interprets a word or phrase as something to be elucidated in the following TOTM	***The notion of disadvantage is very important.*** *The use of German in negotiation between a Stuttgart firm and a Copenhagen firm may be efficient and perfectly logical...*
Prospected	When the prospection in the previous TOTM is fulfilled in the present TOTM			*The notion of disadvantage is very important.* ***The use of German in negotiation between a Stuttgart firm and a Copenhagen firm may be efficient and perfectly logical...***
Verbal Echo	When the TOTM reuses part of the previous sentence in order to change topic but maintain superficial cohesion			*...it is usually easy to agree on a common language for a given discussion with neither party perceiving itself to be at a disadvantage.* *The notion of* ***perceived disadvantage*** *is very important.*
Overlay	When the TOTM appears to be a virtual paraphrase of the previous sentence			*Successful businesses, by contrast, have always been sensitive to the need to respect the language capabilities and preferences of their customers.* ***A Finnish manufacturer would not dream of using Finnish to market a product in Germany or France...***

Internal: When any of the above occurs within the sentence.
Selective reference: Cases of deictic encapsulation and verbal echo, which refer to only part of the previous sentence.
Qualified assignment: When the relationship which the cohesive reference refers to is less than clear.
Double acts of reference: When a sentence contains two acts of reference.

Subsequent adaptations of the analytic system have attempted to classify more implicit cases of encapsulation. Álvarez de Mon y Rego (2001), for example, includes implicit encapsulation in her categorization of encapsulation (see Table 2.10), i.e. where the noun implies the pre-existence of information previously in the text. For instance, in Example 2.13, the noun phrase *the answer* presupposes the existence of a question.

Example 2.13

What is the smallest amount of electric charge that can sit on the head of a pin? ***The answer*** may surprise.

(Cited in Álvarez de Mon y Rego (2001:90)

Table 2.10: Linguistic realizations of encapsulation (Álvarez de Mon y Rego 2001)

Encapsulation category	Description	Examples
By means of nouns	Through the use of textual nouns – the equivalent of Francis' retrospective labels or anaphoric nouns	*This problem, difference, task, case etc.*
By means of nominalization	Whereby a verbal process, attribute or a circumstance is turned into an entity which is participating in another process	*…causes the spray* ***to expand*** *supersonically. Collisions that take place during* ***the expansion*** *cool the clusters*
Encapsulating pronouns	Generally demonstrative pronouns	*this*, *that*
By means of adverbs	Generally circumstance sentence adverbials	*here*, *thus*, *so*, *then*
Implicit encapsulation	Whereby certain vocabulary items imply the existence of	*the answer* presupposes the existence of a question; a

	previous information in the text	comparative or qualifier pre-supposes a previous point, e.g. *a more effective procedure; equally important*

Similarly, Moreno (2004) proposes a three-tier categorization of explicitness of deictic reference in retrospective labels (see Table 2.11). As well as explicit encapsulation, Moreno borrows the term *fuzzy* label from Francis (1994) to refer to occasions where the exact stretch of discourse which is being referred to is not clear. In contrast to Álvarez de Mon y Rego above, implicit encapsulation, which is defined by Moreno as encapsulation in which the reader is left to decide which stretch of the previous discourse is being encapsulated, contains conjuncts. Thus what Sinclair calls *logical encapsulation* is equated by Moreno with *implicit encapsulation*.

Table 2.11: Levels of explicitness in retrospective labels (Moreno 2004)

Retrospective Labels	**Types**	**Examples**
Explicit	Deictic acts expressed in unspecific nominal groups	*These circumstances., the model; the result showing; as*
Fuzzy	Pro-forms	*This means; all this suggests...*
Implicit	Conjuncts; integrated metatextual expressions	*Because; as a consequence; we can conclude that; the main implication is that…*

2.5.3 Prospection

Sinclair proposes three subtypes of prospection: *attribution*; *advance labelling*; and *topic introduction*. These contain some similarities with the categories of prediction in Tadros'

model in section 2.2.4. *Attribution*, for example, would seem to be the equivalent of Tadros' *reporting* category. In using the term *advance labelling* Sinclair employs the same term as Tadros. However, as Álvarez de Mon y Rego (2001) points out, Sinclair's category in this respect is much less specific than Tadros' use of the term. Indeed, it may be argued that such a category is less a *prediction* and more like an *anticipation*, in Sinclair's terms (see Table 2.1), in that it does not mean that the writer has made a firm commitment. It is left unanswered whether Sinclair would exclude the other of Tadros' categories seen in Table 2.2 in the analytic system for written discourse structure or whether it is that they simply did not occur in the example text.

The third subtype of prospection is *topic introduction* based, nominally at least, on the *topic framework* model as proposed by Hazadiah (1993). This model was proposed for spoken discourse and constitutes a series of exchanges between different speakers following a sequence. It is unclear if Sinclair envisages such a sequence to be transferred to written discourse or how. Instead, *topic introduction* is described as simply being when the reader can predict that the text will continue developing the new topic. As Warren (2006:166) notes, the use of the concept of topic in the description of structure as well the process of deciding on what constitutes a 'new topic' would seem to remain intrinsically vague.

2.5.4 Verbal Echo

Verbal echo is described by Sinclair as one of three 'exceptions' to syntagmatic mechanisms described thus far. Verbal echo is a retrospective mechanism, like

encapsulation, but one which does not contain an obvious act of reference. Instead, the term refers to occasions when the text of the moment reuses part of the previous sentence. In Example 2.14, for instance, the noun phrase *perceived disadvantage* in Sentence 13.1 echoes parts of the noun post modifier *perceiving itself to be at a disadvantage* which appears in Sentence 12.4.

Example 2.14

(12.4) ...it is usually easy to agree on a common language for a given discussion with neither party perceiving itself to be at a disadvantage.

(13.1) The notion of **perceived disadvantage** is very important.

(Cited in Sinclair 1993/2004e:99)

In including this type of feature in the analytic system, Sinclair admits that the assertion that there is 'complete textual erasure' as the previous sentence passes to the autonomous plane may, in fact, be a simplification of a much less tidy dynamic process. Instead, he argues that although '[a]wareness of previous words and phrases will die away sharply…the traces, especially of something striking and memorable, may be retained with sufficient clarity to be reactivated' (Sinclair 1993/2004e:91).

For the analyst, however, the admission of such a phrase-to-phrase repeat mechanism is problematic in terms of delimiting the scope of the analytic system. As noted by Mauranen (2012) as regards spoken discourse, a *repeat* (the rough equivalent of a *verbal echo*) is very often not a *verbatim* repetition. Rather, it often entails a varying degree of adaptation from the original s/m with consequent difficulties in the identification of its

occurrence. This difficulty is evident in Example 2.14. Secondly, the distinction between *verbal echo*, a repetition mechanism of structural significance, and a *repetition* as seen in descriptions of lexical cohesion (e.g. Halliday and Hasan 1976; Paltridge 2006; Tanskanen 2006), which, according to Sinclair, does not play a structurally significant role, remains unclear.

2.5.5 Overlay

Another exceptional mechanism which does not contain any obvious act of reference is *overlay*. This describes a situation where two sentences in sequence seem to be almost a paraphrase of each other thus maintaining coherence. It can be conceived as the equivalent in a dynamic description of discourse for which has previously been described as a *matching* relation (Winter 1977, 1994; Sinclair and Mauranen 2006) or as reformulation in spoken dialogue (Muntigl 2007). However, perceiving such virtual paraphrases is, as Sinclair (1993/2004e:92) explains, 'a complex act of interpretation', for which a variety of linguistic comparisons may come into play. For instance, in Example 2.15, Sentence 8.3 is categorized as representing an overlay of Sentence 8.2:

Example 2.15

(8.2) And 'every aspect' most certainly includes language skills – not just for the sales force, but for middle management to keep in touch with the trends in other countries, and for personnel involved in research and development so that they can learn as rapidly and accurately as possible of new ideas and processes. (8.3) The Japanese use Western languages not merely to market their goods, but to improve their products by studying those rivals.

(Cited in Sinclair 1993/2004e:99)

As can be seen in Table 2.12, an overlay can entail the recognition of parallels in the overall structure of the two sentences (e.g. in the parallel structures of the two sentences signalled by the same contrastive connective *but* in d.); of the repetition of certain words, of synonyms and near synonyms (e.g. in b. *not merely to* and *not just*), and, as Sinclair argues, of a similarity in the propositions. Just as we have seen in *verbal echo* in section 2.5.4, the inclusion of overlay and the use of features such as synonymy and antonymy in its identification runs the risk of overwhelming the analytic system with what is normally considered to be instances of lexical cohesion (Paltridge 2006). In addition, as Mauranen (2012) warns, because overlays lack the clear overt textual references of an encapsulation, their identification, particularly if based on semantic similarity, may be inexact.

Table 2.12: Overlay analysis (Sinclair 1993/2004e)

	Sentence 8.3	**Sentence 8.2**
a.	The Japanese use Western languages	language skills
b.	not merely to	not just
c.	market their goods,	for the sales force,
d.	but	but
e.	to improve their products	research and development new ideas and processes.
f.	by studying	so that they can learn
g.	those of their rivals	to keep in touch with the trends in other countries

In sum, the analytic system for written discourse structure provides a realization of a syntagmatic-oriented dynamic description of discourse. However, as seen above, each of the four syntagmatic mechanisms presented by Sinclair presents challenges to the analyst in terms of its exact delimitation.

2.6 Chapter summary

This chapter has described the three models which play a central role in the development of the model employed in this study: LUG; the outline model for spoken discourse; and the analytic system for written discourse structure. As has been seen, they have in common an endeavour to present the language both from a dynamic perspective and in a way that attempts to reflect how users, particularly readers and listeners, experience it, i.e. in a linear fashion. It has also been argued that the phenomena described in these models are essentially the same syntagmatic mechanisms – prospection and completion in the case of grammar and prospection, completion and encapsulation in the case of discourse. This provides this study with the basis, at least, of a model that can integrate the linear description of grammar with that of discourse.

It should also be noted that discourse reflexive elements in the text have been seen to play a prominent role at various junctures in the models, although these have not been described in any systematic way in this chapter. In the next chapter, the literature related to discourse reflexivity and metadiscourse will be reviewed and related back to the linear descriptions presented in this chapter.

CHAPTER 3

LITERATURE REVIEW: DISCOURSE REFLEXIVITY AND METADISCOURSE

3.1 Overview

In Chapter 2, issues relating to the description of linear structure were discussed in both grammar and discourse. In particular, the centrality of both syntagmatic relations and a dynamic perspective in Sinclair's conception of a linear description were highlighted. It has also been noticed that what were seemingly instances of discourse reflexive language coincided with certain categorizations in the linear description of discourse. In particular, discourse reflexive language apparently occurs in: the categories of self-reference, discourse labelling and participant intervention; in certain types of encapsulation, including challenges, especially those that feature a metalinguistic labelling noun; and in certain types of prospection. Nevertheless, the exact relationship between the two systems of description remains elusive at this stage and patently requires further exploration. The aim of this chapter, therefore, is to provide a review of research into the concept of *discourse reflexivity* and the closely related concept of *metadiscourse*, with the particular focus on how the conception of *discourse reflexivity* and *metadiscourse* relates to a linear description of discourse.

In this chapter, then, issues related to the definitions of metadiscourse and discourse reflexivity will be examined in section 3.2. The rest of the chapter will be dedicated to the description and evaluation of two separate models of metadiscourse and discourse

reflexivity, to use the terms proposed by Ädel (2010:70): the *interactive* model of metadiscourse; and the *reflexive* model of metadiscourse in sections 3.3 and 3.4 respectively. The chapter will conclude by asserting that the reflexive model is both the most theoretically sound and more appropriate to employ in conjunction with a linear description of dialogic and polylogic discourse.

3.2 Metadiscourse and discourse reflexivity

There are two terms of central importance to this chapter: *metadiscourse* and *discourse reflexivity*. They both have at their core the central notion that they represent 'discourse about discourse' (e.g. Williams 1981b; Vande Kopple 1985; Mauranen 2007).[12] The term *metadiscourse* was coined in the 1950s by Harris (1970:464), who describes metadiscourse as being discourse which is apart from the 'main material'. However, the term as it is commonly conceived today has its roots in composition theory and academic style guides from the USA in the early 1980s (e.g. Williams 1981a, 1981b; Crismore 1983; Vande Kopple 1985). The term *discourse reflexivity*, on the other hand, is a reasonably new term which has only been used by Mauranen (2001, 2007, 2010, 2012). However, the term *reflectiveness*, referring to the capacity that human language has to refer to itself, was first used by Hockett (1963) as one of sixteen design features of language, with *reflectiveness* being one of the features that sets human language apart from primate language. The concept was introduced to the field of semantics as *reflexivity* by Lyons (1977:5) and to the wider study of linguistics and pragmatics by

[12] There is also a wide variety of related terms in this field, including *metatext* (e.g. Mauranen 1993b); *metalanguage* (e.g. Jakobson 1980), *metatalk* (Schiffrin 1980); *metapragmatics* (e.g. Lucy 1993; Bublitz and Hübler 2007) *metadiscussions* (Crystal 2001) and so on. These may overlap with that of *metadiscourse* and *discourse reflexivity* and may have a different meaning depending on the theorist. They will be explained as and when necessary.

Lucy (1993). The term *text reflexivity* was first introduced to text linguistics by Mauranen (1993a) (see section 3.4.2).

The study of metadiscourse can be divided into two main branches: the 'interactive' model of metadiscourse (also known as the 'integrative model' or the 'broad approach' (Mauranen 1993a, 2012)); and the 'reflexive' model of metadiscourse (also the 'non-integrative model' or the 'narrow approach'). [13] From its inception as a model designed for the description of written academic discourse (Crismore 1983, 1989; Vande Kopple 1985) the interactive model of metadiscourse went beyond the narrow literal conception of what 'discourse about discourse' is, to one which included a wide range of textual features. Hence, the concept of metadiscourse in this model is defined as being all of those parts of written discourse which are considered to be non-propositional (e.g. Vande Kopple 1985; Hyland 1998b) or, alternatively, those parts of the written text which are considered to have an interactive aspect (e.g. Hyland and Tse 2004). It therefore places at the heart of the concept of metadiscourse a view of discourse as 'a social and communicative engagement' (Hyland 2000:109), whereby even monologic written language is seen as being essentially interactive. As will be seen below, this has led to certain metadiscourse theorists (Hyland and Tse 2004; Hyland 2005) employing Sinclair's planes of discourse as a theoretical framework for the interactive model of metadiscourse.

[13] Other theoretical traditions have been employed in alternative models of metadiscourse, such as Speech Act Theory (Beauvais 1989); Rhetoric (Crismore 1989; Nash 1992; Dahl 2004); Relevance Theory (Ifantidou 2005) and so on. Due to space limitations these will not be discussed in detail.

The interactive model of metadiscourse has attracted considerable attention in the last thirty years or so particularly in the study of academic writing (e.g. Williams 1981a; Vande Kopple 1985; Crismore 1989; Mao 1993; Barton 1995; Cheng and Steffensen 1996; Hyland 1998b, 2000, 2005, 2010; Crawford Camiciottoli 2003; Hyland and Tse 2004; Dahl 2004; Abdollahzadeh 2010; Li and Wharton 2012) as well as spoken academic discourse (Heino et al 2002; Thompson 2003; Alessi 2005; Crawford Camiciottoli 2007). It has also been employed less frequently in the analysis in non-academic contexts, such as CEO reports (Hyland 1998d); advertisements (Fuertes-Olivera et al 2001); newspaper articles (Dafouz-Milne 2008); newspaper and magazine editorials (Le 2004; Khabbazi-Oskouei 2013); and wikis (Kuteeva 2011). Such studies have provided insights in a variety of areas, for example: in the comparison of the use of metadiscourse features in different languages (Valero-Garces 1996); in how non-native learners of English differ from their native speaker counterparts in their use of metadiscourse (e.g. Intaraprawat and Steffensen 1995; Cheng and Steffensen 1996; Li and Wharton 2012); in the investigation of the variation in its use in different discourse communities, such as differences in different academic disciplines (e.g. Hyland 1998b, 2000); how the text-type and audience affects its use (e.g. Hyland 1999, 2000) and its effect on the comprehension of written academic texts (Crawford Camiciottoli 2003).

The second model, the reflexive model of metadiscourse, places two related concepts at its centre.[14] The first of these is the concept of reflexivity as described above. A second

[14] Theorists who employ the reflexive model, Ädel (2001, 2005, 2006, 2008, 2010) and Pérez-Llantada (2010) use the term *metadiscourse*. Mauranen uses the term *text reflexivity* (Mauranen, 1993a) and *discourse reflexivity* (Mauranen 2007) specifically in certain publications but also uses *discourse reflexivity*

theoretical basis for this model suggested by Ilie (2003), Ädel (2005, 2006) and Mauranen (2012) is that of Jakobson's (1990) six functions of the speech event and in particular the *metalingual function*. According to Jakobson (1980:86), the term *metalanguage* (from which the term *metalingual* function derives) refers to the level of language 'in which we speak about the verbal code itself' and contrasts with the level of *object language* (Lyons 1977), in which we speak about 'items extraneous to language', i.e. the real world outside the text. *Metalanguage* may be employed to increase the accessibility of the code for the address 'whenever the addresser and/or the addressee need to check up whether they use the same code' (Jakobson 1980:86), i.e. when it is perceived that communication has broken down. In order to do so, Jakobson envisages three ways of performing metalinguistic operations: paraphrasing, providing synonyms, and the making ellipsis explicit.

The defining criterion for the reflexive model of metadiscourse is that of explicitness, i.e. that the text *explicitly* refers to the present discourse (e.g. Mauranen 1993a; Ädel 2006). In this model, then, whether the piece of language is propositional or non-propositional plays no part in considering whether or not the piece of language is considered to be reflexive.

The reflexive model of metadiscourse has seemingly exclusively focused on academic discourse, both written discourse (e.g. Mauranen 1993a; Ädel 2006, 2008; Pérez-Llantada 2010) and increasingly in the analysis of academic spoken discourse, in general

and *metadiscourse* interchangeably in others (Mauranen 2001, 2010, 2012). In this chapter the terms used in the original publication will be used as much as possible.

(Mauranen 2002a, 2002b, 2007); as well as specific spoken discourse text-types: university tutorials (Mauranen 2003); university lectures (Ädel 2010) and seminar debates (Mauranen 2012). It has been employed as a means of contrasting rhetoric by comparing writing in different languages (Mauranen 1993a). It has also been used in the comparison of native speaker and learner English (Ädel 2001, 2006, 2008; Pérez-Llantada 2010); the comparison of spoken and written academic discourse (Ädel 2010) and in exploring the nature of English as a Lingua Franca (Mauranen 2007, 2010, 2012) in an academic context.

In the next section, the characteristics of the interactive model of metadiscourse will be discussed.

3.3 The interactive model of metadiscourse

3.3.1 Metadiscourse as non-propositional

One of the key tenets of early works and subsequent interactive models of metadiscourse is that metadiscourse is non-propositional (e.g. Vande Kopple 1985; Crismore 1989; Crismore and Farnsworth 1990; Crismore et al 1993; Intaraprawat and Steffensen 1995; Cheng and Steffensen 1996; Hyland 1998b). In this connection, Vande Kopple provides the following definition of metadiscourse:

> On one level we supply information about the subject of our text. On this level we expand propositional content. On the other level, the level of metadiscourse, we do not add propositional material but help our readers organize, classify, interpret, evaluate and

> react to such material. Metadiscourse, therefore, is discourse about discourse or communication about communication (Vande Kopple 1985:83).

Central to the early conception of metadiscourse, then, is a model of communication in which two 'levels' or 'planes' (e.g. Crismore 1989) function. The first level, the 'primary' propositional level consists of 'propositions and referential meanings' (Williams 1981a:47). The secondary level is the metadiscourse level, where propositional material is not added to but rather a variety of devices are used to help readers organize, classify, interpret and evaluate the materials in the primary plane. Table 3.1 illustrates Vande Kopple's model based on this premise.

Despite its wide use and acceptance as a model of analysis as seen in section 3.2, the interactive model of metadiscourse has been criticized on several counts. From a linear perspective, Sinclair (2005) criticizes the concept of metadiscourse being on a different 'level' as described above as distorting the way in which users experience language, i.e. linearly, as described in Chapter 2. Another common criticism of the interactive model is the problems evident in delimiting what has been called a 'fuzzy' concept (Crismore 1989; Mauranen 1993b; Hyland 1998b; Hyland and Tse 2004; Ifantidou 2005; Skulstad 2005; Ädel 2006). It has been argued (Hyland and Tse 2004; Hyland 2005; and Ädel 2006) that this has, for the most part, been due to a lack of theoretical rigour in its initial conception. As noted above, the interactive model of metadiscourse can be traced to composition theory and academic style guides where the term metadiscourse was popularized. The first attempts at comprehensive models, namely Crismore (1983) and Vande Kopple (1985), were essentially developed from combining Williams' (1981a,

1981b) categories of metadiscourse with Lautamatti's (1978) categories of *non-topical materials*. Additionally, primary sources of these categories included a wide range of diverse theoretical frameworks, including speech act theory (Austin 1962; Searle 1969), signalling (Meyer 1975) and cohesion (Halliday and Hasan 1976).

Table 3.1: Vande Kopple's (1985) model of metadiscourse

Textual Metadiscourse			
Category	**Function**	**Subcategory & function**	**Examples**
1. Text connectives	To guide readers as smoothly as possible through texts and to help them construct appropriate representations of them in memory	a. **Sequencers**	*First, next, in the third place*
		b. Indicators of **logical or temporal relationships**	*However, nevertheless, as a consequence, at the same time*
		c. **Reminders** about material presented earlier and **statements of material appearing later** in texts	*As noted in Chapter One, as shall be seen in the next section…*
		d. Material the writer is **on the verge of presenting**	*What I wish to do now is develop the idea...*
		e. **Topicalisers:** words that focus attention on a particular phrase, paragraph or whole section' when reintroducing information	*Let us now turn to…*

2. Code Glosses	To help the readers grasp the appropriate meanings of elements in texts by defining and explaining		*For instance*, *defined as*
3. Narrators	To let the readers know who said what		*According to*, *X announced that*
Interpersonal Metadiscourse			
1. Illocution markers	To make speech or discourse act performed explicit		*I hypothesize, to sum up, we claim that, I promise to, for example…*
2. Validity markers	To indicate how the writer assesses the probability of truth of the propositional material	a. **Hedges**: express doubt or 'small notes of civilized diffidences'	*perhaps, may*
		b. **Emphatics**: underscore what the writer really believes or would like the reader to think s/he believes	*Clearly , undoubtedly*
		c. **Attributors**: lead the readers to judge validity of propositional content	*According to X*
3. Attitude markers	To express the writer's attitudes to the propositional material s/he is presenting		*surprisingly, I find it interesting that, it is alarming to note that*
4. Commentary	When the writer addresses the reader directly		*You will certainly agree that*; *you might want to read*

Not surprisingly, the resulting model incorporates such a diverse range of items that it is often difficult to conceive them as forming one singular concept. As can be seen in

Tables 3.1 and 3.2, the inclusion of such diverse elements as logical connectives, modal verbs, the first person pronoun *I*, punctuation marks (e.g. parenthesis and 'scare' quotes), numerals, whole sentences, paragraphs and even whole texts means that the model is in danger of being so wide that it is 'beyond any useful descriptive role' (Hyland 2005:31). As Markkanen et al (1993:143) comment, metadiscourse 'covers such a wide area of language use that … saying that some item in a text is metadiscourse does not say much.' Placing the interactive model in a historical perspective, Mauranen (2012) comments that the interactive model of metadiscourse is a manifestation of a general surge in interest in interactional aspects of discourse from the 1980s, but as a concept, it has now served its purpose and more specific concepts are now required in order to gain more precise insights.

Neither is it clear what exactly is meant by a proposition in these models as it is often used interchangeably with other terms such as *subject*, *topic* and *content*. According to Halliday (1994:70), a proposition is something that can be affirmed, denied, doubted, contradicted, insisted on, regretted and so on, i.e. that it is truth-conditional. Metadiscourse, as non-propositional material, is therefore, seen as parts of the text that cannot be asserted or denied etc. as it is concerned simply with the aiding of the organization of the propositional content of the text or in expressing propositional attitudes.

Ädel (2006:210) demonstrates that the test of truth and falsehood does not apply to such metadiscursive sentences as those seen in Examples 3.1 and 3.2.

Example 3.1

I have discussed X and Y in Chapter 4.

Example 3.2

The topic of the essay is X.

The truth or otherwise of these statements can quite easily be corroborated by referring to the rest of the text. Ädel (2005) therefore argues that the world of the ongoing text may be a specialized world but it is still within the world and so subject to the same conditions of truth etc. that discourse outside the text is. Thus, the assertion that metadiscourse is non-propositional is not tenable in these cases. In LUG terms, this means that discourse reflexivity/metadiscourse can occur in all element types, including M elements, as can be seen in Example 3.3 (the LUG analysis of Example 3.1), where the present discourse is referred to explicitly in lines 1 and 4.

Example 3.3

L	Element	Element type
1	**I have discussed X**	**M**
2	and	OT
3	Y	MS
4	**in Chapter 4**	**MS**

3.3.2 The SFG-inspired model

The diverse textual features of the model, as seen in Table 3.1 are placed by Vande Kopple (1985) under the theoretical framework of Systemic-Functional Grammar (SFG), specifically relating them to two of the three macro-functions as proposed by Halliday (1973) leading Ädel (2005, 2006) to name the model the *SFG-inspired model*.

However, certain theorists (e.g. Mauranen 1993a; Hyland and Tse 2004; Hyland 2005; Ädel 2006) have expressed doubts that the concept of macro-functions can simply be imposed onto the concept of metadiscourse. *Macro-functions* (also called *metafunction*, e.g. Halliday 1994) of language 'are the highly abstract linguistic reflexes of the multiplicity of social uses of language' (Halliday 1973:36) of which there are three: the ideational function; the interpersonal function; and the textual function.[15] The ideational function is a referential, representational or informational function. Elements within it 'are concerned with the content of language, its function as a means of the expression of our experience, both in the external world and in the inner world of our own consciousness' (Halliday 1973:66). As has been seen above, several metadiscourse theorists have endeavoured to equate this function with, what they have termed the 'primary plane' of discourse i.e. the propositional aspect of communication. The other two macro-functions have been seen to be the elements that constitute metadiscourse.

The interpersonal function is concerned with 'language as the mediator role including all that may be understood to express social and personal relations, by the expression of our own personalities and personal feelings on the one hand, and forms of interaction and social interplay with other participants in the communication situation on the other hand' (Halliday 1973:66). This appears in the category of interpersonal metadiscourse, defined as those elements which 'alert readers to the author's perspective towards the propositional information and the readers themselves' (Hyland 1998b:443). As Ädel

[15] The ideational function is later divided into the experiential and logical functions (Halliday 1994:179) the latter functioning at the word group level and so plays no part in metadiscourse theory.

(2006:16) points out, Halliday's interpersonal macro-function includes a broad range of phenomena from modality and mood, to words with negative/positive connotations and intonation, i.e. any linguistic means to express and maintain social and personal relations. As such, many of the linguistic features described by Halliday are generally not included in the SFG-inspired models of metadiscourse, though the criteria for their inclusion or exclusion are rarely made explicit.

The third macro-function is the textual function, which has:

> an enabling function that of creating text, which is language in operation as distinct from strings of words or isolated sentences and clauses. It is the component that enables the speaker to organize what he is saying in such a way that it makes sense in context and fulfils its function as a message (Halliday 1973:58).

Halliday's macro-function again represents a wide range of phenomena, including elements which are not included in textual metadiscourse, such as *theme*, *voice*, *deixis*, *reference* and *substitution*. Yet, textual metadiscourse is limited to:

> Devices which allow the recovery of the writer's intention by explicitly establishing preferred interpretations of prepositional meanings…[which]… help form a convincing and coherent text by relating individual propositions to each other and to other texts (Hyland 1998b:442) .

This would therefore seem to be much narrower than Halliday's conception of the textual macro-function.

It is clear, then, that the relationship between metadiscourse and the textual and interpersonal macro-functions is not a one-to-one relationship. Given the rather dubious nature of the use of Halliday's macro-functions as the theoretical underpinning of metadiscourse, it is unsurprising that theorists have sought alternative theoretical justifications for the model as will be seen in the next section.

3.3.3 The interactive plane-inspired model

Within the group of theorists employing the interactive model of metadiscourse, Hyland's shift from a model based on Halliday's macro-functions (e.g. Hyland 1998b) to one which uses Sinclair's planes of discourse as its basis (Hyland and Tse 2004; Hyland 2005) is the most noteworthy. In this connection, Hyland (2005:37) provides the following planes of discourse-inspired definition:

> Metadiscourse is the cover term for the self-reflective expressions used to negotiate interactional meanings in a text, assisting the writer (or speaker) to express a viewpoint and engage with readers as members of a particular community.

Hyland (2005:38) bases this definition on three key principles: that metadiscourse is distinct from propositional aspects of discourse (as already seen in the previous interactive models in section 3.3.1); that it refers to aspects of text that embody writer-reader interactions; and that it refers only to relations internal to discourse. Hyland

therefore explicitly equates metadiscourse to language on the interactive plane (see section 2.2.4) within which he divides the language into two functions: interactive and interactional, following Thompson and Thetela (1995) and Thompson (2001). Interactive resources, it is argued, are used by the writer to signal the organization of the text, with such decisions being made based on the readers likely understanding and knowledge. Interactional resources, on the other hand, constitute the writer's explicit interventions into the text to comment on and evaluate material (see Table 3.2).

Table 3.2: Interactive plane model of metadiscourse (Hyland 2005)

Category	**Function**	**Examples**
Interactive Metadiscourse	**Help to guide the reader through the text**	**Resources**
Transition markers	Mainly conjunctions and adverbial phrases which help readers interpret pragmatic connections between steps in an argument	*in addition*, *but*, *thus*, *and*, *alternatively*, *although*, *conversely*, *hence*, *rather*
Frame markers	Signal text boundaries or elements of schematic text structure: a. sequencing; b. label stages; c. announce goals; d. shift topic.	 *a. finally*, *to start with*, *firstly* *b. to conclude*, *in short* *c. my purpose is*, *aim* *d. back to*, *so*, *now*, *turn to*
Endophoric markers	Expressions which refer to other parts of the text	*noted above*, *see Fig.*, *in section 2*
Evidentials	Metalinguistic representations of an idea from another source, which guide the reader's interpretation and establish an authorial command of the subject	*according to X*, *Z states*

Code glosses	Parts of text which supply additional information by rephrasing, explaining or elaborating what has been said to ensure the reader is able to recover the writer's intended meaning	*Namely*, *e.g.*, *such as*, *in other words*, *( )*
Interactional Metadiscourse	**Involve the reader in the text**	**Resources**
Hedges	Devices which indicate the writer's decision to recognize alternative voices and viewpoints and so withhold complete commitment to a proposition.	*Might*, *perhaps*, *possible*, *about*
Boosters	Devices which allow writers to close down alternatives, head off conflicting views and express their certainty in what they say.	*in fact*, *definitely*, *it is clear that*
Attitude markers	Devices which indicate the writer's affective rather than epistemic attitude to propositions – surprise, agreement, importance, obligation, frustration etc.	*unfortunately*, *I agree*, *surprisingly*, *amazed*, *curiously*, *dramatically*, *surprising*, *unbelievably*
Self mention	Devices which refer to the degree of explicit author presence in the text.	*I*, *we*, *my*, *me*, *our*
Engagement markers	Devices which explicitly address either to focus their attention or include them as discourse participants.	*consider*, *note*, *you can see that*, *?*, *by the way*, *follow*, *go to*, *let's*, *need to*, *should*, *you*

It should be noted that certain aspects of what Sinclair envisages as featuring prominently on the interactive plane do feature in Hyland's model in Table 3.2. The subcategory of *announce goals* in *frame markers* would seem to coincide with what Sinclair calls *self-reference* (see section 2.2.4), e.g. *In this chapter we introduce…*, and may also contain

some sort of prospection, that meets Tadros' (1994) conditions for prediction, seen, for instance in the stem: *There are several reasons why*. Similarly, the category endophoric markers would seem to coincide with what Sinclair calls *cross-referencing*. In addition, the category of *engagement markers* would also seem to provide the possibility of including what Sinclair and Mauranen term OI elements in LUG, i.e. those elements which function on the interactive plane by providing explicit means of the management of the interaction.

Furthermore, following Mauranen (1993a), Hyland refines the conception of the textual function of language such that it is only those conjunctions and adverbial phrases with 'internal relation', i.e. those related to connecting steps of an exposition, which should be considered to be functioning on the interactive plane and not those which demonstrate an 'external relation', i.e. those related to connecting activities in the outside world on the propositional or autonomous plane (see section 2.3.4). This distinction is one made by Halliday (1994:325):

> Many temporal conjunctives have an 'internal' as well as an 'external' interpretation; that is, the time they refer to is the temporal unfolding of the discourse itself; not the temporal sequence of the processes referred to.

This is most clearly seen in time sequencers, where an internal interpretation of *firstly*, *secondly, lastly* etc. is related to the organization of the text whereas the external interpretation is the temporal sequencing of events. Those with internal relations are

included as metadiscourse in Hyland's model as functioning on the interactive plane whereas those that show external relations function by organizing propositional information and so are seen to be acting on the autonomous plane.

However, despite these refinements, Hyland stops short of implementing wholesale changes to his existing SFG-inspired model (Hyland 1998b). In fact, as can be seen in Table 3.2, despite this radical shift in theoretical justification of the concept of metadiscourse, the model itself remains largely unchanged. This would seem to be the equivalent of replacing the ground floor of a building yet at the same time keeping the upper floors intact.

It is also clear that Hyland's conception of the interactive plane, as manifested in the model, does not reflect Sinclair's dynamic view of discourse or the analytic system for the description of text structure (Sinclair 1993/2004e), which, as we have seen in section 2.5, is Sinclair's most coherent application of the functioning of the interactive plane in written texts. In particular, it would seem that Hyland's model does not consider the most crucial aspect of Sinclair's model of discourse, that it is a dynamic model moving from one text of the moment to the next (see section 2.2.4) and that each text of the moment functions on the interactive and autonomous plane simultaneously. As Hunston (2000:183) comments, 'some sentences draw attention to their status on the interactive plane more *explicitly* than others' (my italics). This difference in interpretation of the interactive plane is most clearly seen in Hyland's omission of cases of encapsulation using metalinguistic labelling nouns (as seen in section 2.2.4), which would seem to be

clear ‘explicit interactional signals’ of activity on the interactive plane. However, as these are considered to be propositional they are excluded from Hyland’s model. Neither is any mention made of the concept of *plane change*, which, as we saw in section 2.2.4, would seem to be an occasion where discourse reflexive language plays a crucial role. Indeed, there is no real sense that Hyland has embraced the notion of a dynamic model of discourse with its ‘continuous movement’ as Sinclair envisages. Instead, Hyland’s interactive plane-inspired model reflects a partial representation of the parts of the text in which the writer explicitly announces his/her presence on the interactive plane through a slightly modified version of Vande Kopple’s original taxonomy.

In relation to the third broad goal of this thesis as outlined in section 1.2.2, it is also difficult to envisage how such an interactive model of metadiscourse (whether it be Vande Kopple’s or Hyland’s) can cope with dialogic or polylogic discourse. In fact, it could be argued that the interactive model of discourse is only saved from being overwhelmed by the amount of interactive elements in discourse because it has traditionally been limited to the description of a type of discourse which has little or no overt interaction. For instance, Example 3.4 is a standard exchange in spoken discourse.

Example 3.4[16]

Why? Did **you** wake up late today**?**	I
Yeah, pretty late.	R
Oh dear.	F

[16] Example 2.10 in Chapter 2.

If we apply the criterion that overt interaction is to be considered to be metadiscourse, it could easily be argued that the entire exchange is overtly functioning on the interactive plane. In terms of Hyland's model, even if we limit it to 'non-propositional material', the following features would be considered to be metadiscourse:

Engagement markers: *Why?*; *you* ; *?; Yeah*

Attitude marker: *Oh dear*

Similarly, Sinclair and Mauranen (2006:115–116) provide an example of a spoken dialogue in which over half of the elements are O elements in LUG terms and would be considered to be metadiscursive following Hyland's categorization, as presumably would other 'non-propositional' interactive elements prevalent in dialogic discourse, such as vocatives (Leech 1999). As Mauranen (2010) comments, now that we have a better understanding of the variety of interactional features in language, be it hedges, vague language or fillers, there seems to be little gained by lumping them all together under one banner, especially if the items grouped under the banner *metadiscourse* have little to do with the original concept of being 'text about text'.

3.4 The reflexive model of metadiscourse

3.4.1 Reflexivity and explicitness

As has been seen already in section 3.2, an alternative to the interactive model of metadiscourse is the reflexive model in which the defining criterion is language's capacity to refer to itself, or more accurately, as Ädel (2006:1) notes, moments where speakers and writers choose to use language to refer to or describe language itself. The

concept can incorporate a diverse range of linguistic phenomena (e.g. Lucy 1993). However, Mauranen argues that the key to *text reflexivity* is that of the explicitness of the reference to the present text or, in other words, the overt demonstration by the author of his/her awareness of the text as text. Thus, text reflexivity can be:

> characterized as interactive use of language about the ongoing text with explicit authorial presence…text reflexivity is not an indication of the writer's presence in the text in general but an explicit expression of a writer's awareness of the current discourse as text (Mauranen 1993a:152).

Similarly, Ädel (2006:66) defines reflexivity as 'ways in which speakers and writers use language to attract attention to aspects of language itself, rather than to ideas, arguments, or facts that are not directly to do with language or linguistic matters.'

Table 3.3 illustrates the model of text reflexivity presented by Mauranen (1993a) in which text reflexivity is divided into high explicitness and low explicitness. Text reflexivity of high explicitness includes those expressions in which the author unambiguously demonstrates awareness of text as product, i.e. the outcome of a process or the writing process itself.

Table 3.3: Model of text reflexivity (Mauranen 1993a)

High Explicitness			
A1.	References to the text	Nominal references; adverbs of place and time	***The paper*** *concludes*; *In this* ***article****…*; *In the following* ***section***; *Let us* ***now*** *explore the implications…*

A2.	Discourse labels	Verbs of illocution or similar expressions	*To illustrate*; *as noted earlier*; *stated formally*; *briefly*; *this argument*; *before proceeding*
A3.	Addressing the reader	Direct addresses to reader as reader or references to reader which are not directed to reader	*Note that…; recall that…* ; *The tedious algebra is left to* ***the reader***
Low Explicitness			
B1.	Internal connectors	Connectors expressing relations between elements of text rather than between propositions or states of affair in outside world	*However*,…; *Second...*; *Also…*
B2.	Discourse labels	Verbs with ambiguous reference between action within present text or outside it	*This question will be discussed; it is reasonable to think*
B3.	References to the text	References not clearly separated from content or augment of text	*It is now evident*; *As a first step*
B4.	Addressing the reader	Less direct addresses to reader or address regarding action outside text	*There is reason to remember*

Text reflexivity of high explicitness, therefore, is expressed in Mauranen's model through explicit references to the text (e.g. *essay*, *paragraph*, *paper*, *sentence*, *section* etc.). In this way issues relating to whether such references are propositional or non-propositional, which has preoccupied the interactive models of metadiscourse (see section 3.3.2) are circumvented. By extension, metalinguistic labelling nouns (Francis 1994) such as *your complaint* or *this claim* (Mauranen 2010:18), which play a central role in the syntagmatic mechanism of encapsulation on the interactive plane in Sinclair's dynamic model of

discourse (see Chapter 2) would be included in this model. The category of discourse labels is represented for the most part by illocutionary verbs. These have been recognized in the majority of models of metadiscourse and metatext (e.g. Lautamatti 1978; Vande Kopple 1985; Beauvais 1989; Crismore and Farnsworth 1990; Crismore et al 1993; Mauranen 1993b; Longo 1994; Valero-Graces 1996; Vande Kopple 1988; Bunton 1999; Rahman 2004; Bondi 2005; Ädel, 2006) as a means of making the speech or discourse act explicit. Mauranen notes that in the third category of addressing the reader, if the reference is to be considered to be of high explicitness, it must contain a clear reference to the reader as reader partaking in the processes of communication and transferring of information through the text rather than references to the reader as an actor in the outside world. Curiously, as Ädel (2006) notes, the writer him/herself appears to be missing from the model, though one may argue that s/he would inevitably be present in the discourse label category, i.e. accompanying a discourse verb.

Text reflexivity of low explicitness, on the other hand, includes expressions which are not unambiguously referent to the present discourse alone, including text connectives with internal relations, as seen in Hyland's model in section 3.3.3. However, certain text connectives with internal relations would seem to be as least as prominent on the autonomous plane as they are on the interactive plane. For example the connective *then* in Example 3.5 signals a logical condition of a supposition in an argument:

Example 3.5

> If you link the swipe card to your mobile number then you can use it at any one of over 60000 TopUp points.
> (Cited in Hyland (2005:46))

In this instance, *then* makes the semantic or logical relation between the two concepts in the outside world explicit, i.e. it makes explicit the link on the autonomous plane. In contrast, it provides a less than explicit reference to its own role on the interactive plane, i.e. as a link between the two parts of the text as text, if compared to a clearly explicit equivalent reference to the text such as: a *consequence of the previous statement*. As we have seen in section 3.3.3, Hyland includes such cases as instances of interactive metadiscourse. Ädel (2006), on the other hand, excludes connectives from her model of metadiscourse as does Mauranen (2010) for not making an explicit enough reference to the discourse.

Hence, whilst the inclusion of categories of text reflexivity of low explicitness augments the external validity of the model in reflecting the reality of the ambiguous nature of real data, it also runs the risk of making the model inexact to the extent that its reliability and viability may be undermined. As will be seen in Chapter 4, a defining criterion of similarly high explicitness for discourse reflexivity as proposed by Mauranen (2010) is applied in this study.

3.4.2 The speech event model

The reflexive model of metadiscourse has produced two recent approaches. The first approach can be termed the *speech event model* (e.g. Ädel 2006, 2008 2010; Pérez-

Llantada 2010), in which Ädel proposes an alternative model to the interactive model seen in Tables 3.1 and 3.2. This model is based on the concept of reflexivity and Jakobson's functions of the speech event as described in section 3.4.1. Ädel defines metadiscourse as:

> reflexive linguistic expressions referring to the evolving discourse itself or its linguistic form, including references to the writer-speaker qua writer-speaker and the (imagined or actual) audience qua audience of the current discourse (Ädel 2010:75).

Ädel therefore envisages metadiscourse as being the interaction of three of Jakobson's functions, the metalinguistic; the expressive and directive functions. This represents the interaction between the text, writer and reader. In any given example of metadiscourse according to Ädel, one of these three will be most prominent with the equivalent text/code, writer and reader-oriented metadiscourse. Ädel emphasizes, however, that the metalinguistic function is indispensable for it to be included as metadiscourse. Following Mauranen (1993a), as seen in section 3.4.1, this means that simply referring to the participants is not enough for it to be considered to be metadiscourse. The writer must be referred to in his/her capacity as a participant in the discourse, not as an actor in the outside world, or in Hunston's (2000) terms as *text- constructor* as opposed to *informer* (see section 2.2.4). This coincides with research into overt authorial presence and the concept of the author taking on several roles in writing; the metalinguistic role coinciding roughly with the categories of *I as guide*, *navigator* or *architect* (Tang and John 1999; Sheldon 2009; Chávez Muñoz 2013) or *self as author* (Herriman 2007).

The development of the model based on Jakobson's functions reiterates two points made by Mauranen in the model of text reflexivity. Firstly, metadiscourse and evaluation are conceived as being conceptually distinct. They should then, according to Ädel, be treated as two distinct categories apart from a few exceptional cases (e.g. the performative verbs *argue* and *claim* have both an evaluative and reflexive function). When no evaluation is involved but the writer is referring to him/herself or the reader as actors in the outside world, Ädel refers to this as 'participation', again conceptually distinct from metadiscourse. A similar distinction is made by Bunton (1999) between writer and researcher in post-graduate theses. Secondly, unlike the interactive model of metadiscourse, discourse about other discourse (*evidentials* in Hyland's terms) is not included as metadiscourse as it is defined as only referring to the present ongoing text. An evidential, on the other hand, has an intertextual reference.

Ädel presents two separate models based on this theoretical grounding, the first (Ädel 2006) for written academic discourse; the second (Ädel 2010) for both written academic discourse and monologic spoken academic discourse.

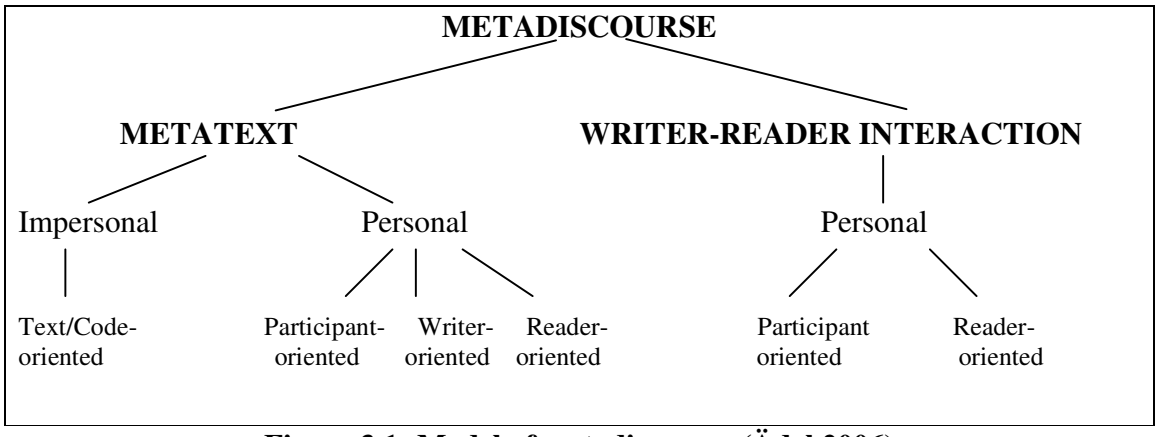


Figure 3.1: Model of metadiscourse (Ädel 2006)

In Ädel (2006), metadiscourse is divided into two categories: *metatext* and *writer-reader interaction* (see Figure 3.1 and Table 3.4). *Metatext* (see also Mauranen 1993b; Bunton 1999; Dahl 2004; Fløttum et al 2006) refers to the text itself or the writing of it and the writer's and/or the reader's discourse acts. This in turn is divided into impersonal and personal metatext. Impersonal metatext includes elements where the discourse participants are left implicit, whereas personal metadiscourse includes an explicit reference to the writer (i.e. writer-oriented metadiscourse), reader (i.e. reader-oriented metadiscourse) or both (i.e. participant-related metadiscourse). *Writer-reader interaction* focuses on the relations between the writer and the reader e.g. by addressing the reader directly, anticipating the reader's response etc. In Ädel (2010), the taxonomy is limited to personal metatext and *audience interaction*.

Of the two main categories, the category of writer-reader interaction would seem to be less clearly delimited of the two. From the examples provided (Ädel 2006), certain elements in this category are clearly reflexive, such as those that explicitly use the words:

reader as a vocative, or the verb *say*, e.g. *I am going to disappoint you when I say* etc. However, several instances provided do not contain any overtly discourse reflexive elements. For instance, in the category of *Anticipating the Reader's Reaction* the following example is provided:

Example 3.6

Yes, it is true! Terrible, I know!

(cited in Ädel 2006:73)

Ädel argues that *I know* constitutes an example of metadiscourse in this context because the initial *Yes*, indicates that a mock exchange is taking place between the writer and the imagined reader. The phrase *I know*, according to Ädel, also anticipates the reader's response to the previous statement. However, if this is to be considered to be sufficiently explicit to be metadiscourse, the question arises what would be classified as metadiscourse in Example 3.7, from the IMDb corpus, where *I know* is used, not as a means of anticipating the reader's reaction, but as an actual means of reacting to a previous poster:

Example 3.7

A: u're comparing Bruce Willis to Ashton 'i'm awesome' Kutcher! Wot a moron!

B: I know, Ashton is so much better.

If *I know* in this context is metadiscursive, then every response in such an exchange would have to be classified as metadiscursive. It would seem that the same criticisms

levelled at the interactive model of metadiscourse in section 3.3.3 can apply here as well, i.e. that the category of *writer-reader interaction* as defined by Ädel is only meaningful in text-types where there is little or no actual direct interaction between the writer and reader. Ädel's model as it stands is applicable to monologic discourse whether it be written monologue (Ädel 2006) or spoken monologue, such as university lectures (Ädel 2010). However, in a text-type which is dialogic or even polylogic, the categorization as it stands would be untenable as notions of metadiscourse being present if there is a mock exchange would inevitably have to be extended to cases of real exchange.

Table 3.4: Taxonomy of metadiscourse (Ädel 2006)

	Impersonal Metatext	
	Discourse function	**Example**
	References to the text/code where the whole or part of the text is referred to. Reference below level of paragraph called references to the Code.	*the paper, in the following section*
	Phoric markers point to the various portions of the current text, helping the reader to navigate through the text.	*As noted above.*
	Discourse labels indicate the textual function of the part of the text in question.	*Mention, call, such as, outline*
	Code glosses give clues to the proper interpretation of elements, comment on ways of responding to elements in texts or call attention to or identify a style.	*Briefly, bluntly, mean*
	Personal Metatext	
	Discourse function	**Example**
CODE	**Defining** explicitly comments on how to interpret terminology.	*What do we mean by ___ then?* *We have to consider our definition*
	Saying involves general *verba dicendi* such as *say*,	*What I am saying is_*

	talk, or *write, in which the fact that something is being communicated is foregrounded.*	*A question I ask myself is_*
TEXT	**Introducing the Topic** gives explicit proclamations of what the text is going to be about, which facilitates the processing of the subsequent text for the reader.	*In the course of this essay, we shall attempt to analyse whether_* *I will discuss_*
	Focussing refers to a topic that has already been introduced in the text; announces that the topic is in focus again, or it narrows down.	*Now I come to the next idea which I presented at the beginning* *I will only discuss the opponents of_*
	Concluding is used to conclude a topic.	*In conclusion, I would say that_*
	Exemplifying explicitly introduces an example.	*As an example of_, we can look at* *If we take _as an example of_*
	Reminding points backwards in the discourse to something that has been said before.	*As I mentioned earlier,_* *As we have seen,_*
	Adding overtly states that a piece of information or an argument is being added to the existing one(s).	*I would like to add that_*
	Arguing stresses the argumentative discourse act being performed in addition to expressing an opinion or viewpoint. Verbs used as performatives.	*The _which I would argue for is_*
	Contextualising exhibits traces of the production of the text or comments on (the condition of) the situation of writing.	*I have chosen this subject because_* *I could go on much longer, but_*
	Writer-reader interaction Functions	
	Discourse function	**Example**
PARTICIPANT	**Anticipating the Reader's Reaction** pays special attention to predicting the reader's reaction to what is said, e.g. by explicitly attributing statements to the reader as possible objections or counterarguments conceived by him.	*I do realize that all this may sound_* *You probably never heard of_ before either* *You would be very surprised at_*
	Clarifying marks a desire to clarify matters for the	*I am not saying_, I am merely*

	reader; motivated by a wish to avoid misinterpretation. Negative statements are common.	*pointing out that_* *By this I do not mean that_*
	Aligning perspectives takes it for granted that the reader takes the writer's perspective. The reader's agreement is presupposed.	*If we [consider/compare]_, we [can/will] [understand/see] _*
	Imagining Scenarios is a 'picture this' type of encouragement that (often politely) asks the reader to see something from a specific perspective. It allows writers to make examples vivid and pertinent to the reader.	*If you consider_, you can perhaps imagine_* *Think back to when you were_* *When you were that age_*
	Hypothesising about the Reader makes guesses about the reader and his knowledge of attitudes.	*You have probably heard people say that_*
	Appealing to the Reader attempts to influence the reader by emotional appeal. The writer persona conveys her attitude with the aim of correcting or entreating the reader.	*I hope that now the reader has understood_* *In order for _ You and I must keep our minds open*

3.4.3 Discourse reflexivity in dialogic discourse

3.4.3.1 Managing discourse strategically

A second recent approach to the description of metadiscourse/ discourse reflexivity within the reflexive model has been seen in recent work by Mauranen (2001, 2007, 2010, 2012) mainly related to spoken English as a Lingua Franca (ELF). In this research Mauranen has argued that the use of discourse reflexivity in dialogic discourse is distinct from its use in monologic discourse and, as the vast majority of discourse that has been examined in the literature regarding metadiscourse and discourse reflexivity has been monologic written discourse, that these roles have not been appreciated thus far.

Mauranen (2010) comments that discourse reflexivity in monologic texts can, at a fundamental level, be considered to represent an 'act of consideration' for the listener in a spoken text just as for the reader in written discourse. This is seen most obviously in the type of discourse reflexive elements that guide the reader/listener by referring to and commenting on the ongoing discourse. These types of items are present in most models of metadiscourse or text reflexivity and related to the organization of the discourse. They tend to be clausal or longer phrases and generally carry out such functions as: prospecting ahead, making retrospective references, resuming earlier topics and making transitions (Mauranen 2012). They include: *we have been looking at…; here we have a statement which...* etc. and appear, for instance, in the taxonomy of metadiscourse designed for both spoken and written discourse (Ädel 2010) as *previewing*, *reviewing*, *introducing topics* and *endophoric marking*. However, the equivalent of these categories also feature in earlier interactive models of metadiscourse, such as Vande Kopple (1985) where they are categorized as *reminders*, *topicalisers* and *markers of what is about to happen*. In Hyland (2005) these are represented by various types of *frame markers*: *sequencing*, *label stages*, *shift topic* and *endophoric markers*. In fact, what Ädel (2010) shows in comparing the use of personal metadiscourse in academic written articles with its use in university lectures is that the discourse functions carried out by personal metadiscourse in general in written and planned spoken monologues, in an academic context at least, are broadly similar. This coincides with the findings of Mauranen (2012) in the context of ELF.

The motives for the use of such elements may not be quite as straightforward and transparent, however. Metadiscourse has been seen, for example, as a means of

establishing ‘solidarity’ (Fuertes-Olivera et al 2001) between the writer and reader and of ‘galvanizing support’ (Hyland 2005:74). It has also been seen as a manifestation of the norms, conventions, values and ideologies of the discourse community in which the participants are acting (Dillon 1981; Hyland 1998b, 1998c, 2005; Ilie 2003) as well of the relative power relationships within that community (e.g. Hyland 2000). However, in the end, discourse reflexivity in monologic text, be it written or spoken, is seen as a means of limiting interpretations of the reader so that there is an alignment of the reader(s) with the writer’s intentions (Mauranen 1993a; Hyland 2005) and this is done amongst other things through the means of these signposting phrases.

In contrast, according to Mauranen (2012), in dialogic and polylogic spoken discourse reflexivity is much more diverse and multifarious. According to Mauranen (2010), discourse reflexivity in dialogic spoken discourse functions in two broad types of roles. The first of these is similar to those described for monologic discourse above, i.e. by making discourse more explicit or precise. These relate to what can be termed managing the interaction through providing clarification, seeking clarification, setting expectation prospectively and evaluation. The second role can be termed managing discourse strategically and is related to the ongoing negotiation of roles in the discourse as well as the speaker imposing his/her will on the present discourse in order to intervene, allow others to participate and dictate the rules of the discourse such as topic etc. These can be seen in Table 3.5.

Table 3.5: Discourse reflexive functions in dialogic spoken discourse (Mauranen 2010)

	Function	**Example**
A.	**Make discourse more explicit and precise**	
1	Indicating how speech relates to current state of discourse; how the speakers' contributions are to be taken	*I mention that just for those who are interested it is totally irrelevant to what I mean*
2	Indicating how the speaker has understood interlocutor's speech	*Were you saying that or am I just hearing you differently*
3	Prospecting ahead	*Now I just wanna clarify*
4	Retrospectively labelling	*your complaint is that this claim is not the whole truth*
B	**Manage discourse strategically**	
1	Taking the floor	*May I ask a question here*
2	Yielding or offering floor	*I was wondering if you could comment*
3	Imposing order on discourse	*So lemme start what I'm gonna do is I'm gonna talk for a while*
4	Negotiating terms of conversation	*What exactly should we be discussing*
5	Avoiding or shelving topics	*I don't have time to discuss that today*
6	Resuming topics from earlier stages	*Remember when we were talking about*
7	Evaluating interlocutors' contributions	*You guys have brought up a lot of important points*

Such a managerial role of discourse reflexivity overlaps with the concept of metapragmatic utterances, i.e. utterances whose role it is to manage the communicative effectiveness of the dialogue (Tanskanen 2007). In a study of online mailing lists and message board discussions (see also Arendholz 2013), Tanskanen identifies three functions of such metapragmatic utterances: judging the appropriateness of one's own or other's contributions, e.g. *I apologise for the tortured prose*; controlling and planning

interaction, e.g. *I'd like to hear what others think*; and providing feedback on ongoing interaction, e.g. *I'm absolutely loving this dialogue!* According to Tanskanen these utterances are employed as a means of participants overtly collaborating in order to minimize the chance of being misinterpreted or judged to be inappropriate.

Mauranen (2012) indicates, however, that metadiscourse in this regard may be employed for more manipulative purposes. According to Mauranen, metadiscourse is often used when a speaker takes on a chairperson's role in a seminar, seen in phrases such as *if we take now* and *and then go after a while to…* . In this context, as Mauranen (2012:175) comments, 'the chair is not organising his own talk, but that of the entire speech event, with consequences for other speakers' freedom of movement.' In this conception of discourse reflexivity, therefore, discourse reflexivity has moved from being a means of aiding the reader or listener to a means of manipulating the discourse behaviour of the other participants.

Mauranen further argues that there are a number of occasions where discourse reflexivity co-occurs with hedging devices. This is seen in the use of *just* (Mauranen 2001) in phrases such as: *i just wanted to point out*; *i just want to try to get across*; and *i was just saying*. It is also seen frequently in phrases containing *ask* such as *I would like to ask (you)* and *I want(ed) to ask (you)*. Mauranen (2012:170) argues that this is an example of *discourse collocation* i.e. two discourse phenomena co-occurring frequently, whereby the combination of hedging, which opens up negotiation acts, and discourse reflexivity which

represents a reduction in the negotiability of interpretations, acts as means of maintaining the balance of the ongoing discourse.

3.4.3.2 Structural role

As has been discussed above, in monologic discourse, such discourse reflexive elements as reminders, topicalisers and so on serve a function in the linear structure of the monologue in that they prospecting what is going to come up and provide a means of retrospectively linking back to the previous discourse. However, in dialogic discourse as we have seen in section 3.4.3.1, the presence of other active discourse participants means that any such prospection is likely to be employed for controlling and managing the direction of the discourse, just as Sinclair (1992) argues in relation to the exchange structure in the outline model for spoken discourse (see section 2.4). Mauranen (2012) proposes that, in ELF at least, other-oriented metadiscourse (i.e. metadiscourse that refers to the other discourse participants) has three main functions:

1. Elucidation: the speaker wants the previous speaker to clarify, confirm, or expand on what has said e.g. *are you saying*….
2. Interpretation: speaker offers an interpretation e.g. *so you're saying*.
3. Springboard: the speaker paraphrases previous speakers meaning for a point of departure e.g. *I found this really interesting that you mention that* …

Thus, discourse reflexivity acts as a signal as to how the interaction structure is being controlled. In the first category, an *elucidation* prospects and invites the other participant

to participate in an exchange, which, as Mauranen (2012) argues has undergone a plane change and so is occurring on the interactive plane being a clarification of what has been said. In the second category, *interpretation*, the other participant is not invited to participate but the interpretation does not preclude the speaker coming back with a further correction. In the third category, *springboard*, the further intervention in the meantime of the other participant is precluded. Hence, discourse reflexivity can be seen to be playing a central role in how the discourse is being structured and how much freedom or restrictions speakers have to intervene.

3.5 Chapter summary

This chapter has provided an overview of the vast amount of literature which has been dedicated to the subject of metadiscourse and discourse reflexivity over the last thirty years. What is readily apparent, as Ädel and Mauranen (2010) comment, is that apart from the initial definition, theorists agree on little else. However, the aim of this chapter was to examine the previous research specifically in order to find which concept is most theoretically sound and conducive to be employed in conjunction with a linear description of both grammar and discourse in written dialogic discourse. It was found that despite claims that the model is based on Sinclair's concept of the interactive plane, Hyland's (2005) model retained the shortcomings of other similar models. Specifically, the concept of metadiscourse, if based on being interactive, becomes untenable in overtly interactive discourse, as much of the discourse would have to be categorized as such. A similar shortcoming was seen, at least in part, in Ädel's (2006, 2010) models.

It seems clear that the sound and applicable defining feature of discourse reflexivity should be that the reference to the present discourse be explicit. Discourse reflexivity should also include occurrences in all parts of the text whether it be in propositional or non-propositional material or, in the terms of LUG, whether it occurs in M, OI or OT elements. This would allow the inclusion of clausal units in the text that prospect and refer retrospectively to other parts of the text, as we have seen in most models of metadiscourse and discourse reflexivity (e.g. Vande Kopple 1985; Hyland 1998b; Hyland 2005; Ädel 2010; Mauranen 2012) but would also allow the inclusion of instances of explicit encapsulation (see Chapter 2). This in turn provides the opportunity to explore both the role of discourse reflexivity fully in Sinclair's dynamic model of discourse and as a means of managing and controlling the discourse as recently explored by Mauranen (2012). These, then, will be the characteristics applied to the model of analysis presented in the next chapter.

CHAPTER 4

METHODOLOGY AND PROCEDURE

4.1 Overview

In the previous two chapters the literature related to Sinclair and his associates' conception of linear descriptions has been explored. It has been suggested that there is a commonality in these descriptions which may be employed for an integrated description of both grammar and discourse based on the use of syntagmatic mechanisms. It was also shown that a view of discourse reflexivity based on the criterion of explicit reference to the present discourse was both theoretically sound and allows the application of the concept of discourse reflexivity to both monologic and dialogic/polylogic discourse. This chapter presents the system of analysis which attempts to incorporate all of these aspects for use in this study.

This chapter is divided, then, into three main sections. The first section, section 4.2 describes the compilation of the IMDb corpus, which is the corpus employed in this study. In section 4.3, the linear system of analysis developed for this study is presented. The description is divided into the system of analysis for element relations and the system of analysis for linear unit relations. In section 4.4, the definition of discourse reflexivity proposed in Chapter 1 is expanded upon. Finally the procedures for the study will be outlined, specifically the use of the *UAM CorpusTool* (O'Donnell 2009) in the coding of the data and the use of *Wordsmith Tools* (Scott 1996) in calculating the results presented in Chapters 5 and 7.

4.2 IMDb corpus

4.2.1 Compiling a corpus of online message board threads

As stated in Chapter 1, the IMDb corpus is made up of data taken from the thousands of message boards which are available on the IMDb website. These are message boards which are dedicated to specific actors, films, directors, film genres and so on. Those who contribute to these message boards (hereafter *posters*) have to register with the site to be able to contribute. Their personal data is not disclosed. However, as one of the most popular international websites on the Internet, posters are inevitably from different countries and linguistic backgrounds. Specific mention is made, in the IMDb corpus data, to posters living in the USA, England, Scotland, Albania and Japan. The message boards are moderated and a set of *Terms and Conditions* listing types of inappropriate behaviour is posted on the website. Posts are regularly deleted by the moderator although the reasons are not provided.

Previous corpora of online message boards have applied different strategies in regard to compiling a corpus depending on the research purposes. Lewin and Donner (2002), for instance, in examining the frequency of features typically associated with CMC in online message boards, compiled a message board corpus through selecting decontextualized individual messages. This was done by taking every fourth message from five different message boards and from taking a sample from each author only once. In contrast, other researchers (e.g. Collot and Belmore 1996; Claridge 2007; Arendholz 2013) have used the thread as the unit of compilation of message board corpora. This has the advantage of maintaining the integrity of the interaction between posters. As this study aims to study

the role of discourse reflexivity in both interaction between participants and within the individual post the latter approach was followed here in order to provide the context of each message.

As the thread was to be used as the basic unit of compilation, what exactly constitutes a thread had to be decided. As Claridge (2007) points out, a thread always has a clear beginning in that it has a subject line but it has no structurally signalled end as one does, for instance, at the end of a telephone call. Rather, a message board thread simply stops when there is a lack of interest and no-one adds a post. Even then, someone may add a post weeks, months or even years after the final post. It is in effect then dormant as opposed to extinct at that point, meaning that a thread is never definitely complete. In terms of corpus compilation, this means that the corpus can only be said to capture the particular state of the thread at the point of compilation. The thread may or may not have been added to two minutes, two hours or two months after the corpus is compiled. Claridge recommends that in the gathering of message board threads for a corpus, the thread should be no shorter than sixteen messages in order to choose threads which show the development of interaction. This was felt to place an artificial restriction on the random selection of threads as thread of two or even one message may be due to it being relatively recent or unsuccessful in provoking interaction and so threads even with one message were included. Following Claridge, a limit of approximately 1,000 words to be gathered per thread was set in order to reduce the possibility of over-representation by one thread or by a small number of posters. If the thread was longer than 1,000 words then the complete nearest turn to the 1,000 words was included, meaning that the

integrity of the final turn was not compromised. As noted above, a thread does not have a clear structural ending so cutting the thread short could be considered to be the equivalent of sampling up to an earlier date.

Given that the structure of the interaction between posts was likely to be one of the principal focuses of the study, capturing the structure of the whole thread was of paramount importance. The *thread* view option (see Figure 4.1) in the IMDb website shows the subject heading of each post such that it can be seen which previous post the poster is replying to. This is indicated by each message being further indented as seen in Figure 4.1. The *nest* view option shows the same information as the *thread* option but with each post in full as can be seen in Figure 4.2. It is important to note that the *thread* and the *nest* options are not chronological in their ordering but rather they are organized by which post the present post is responding to.

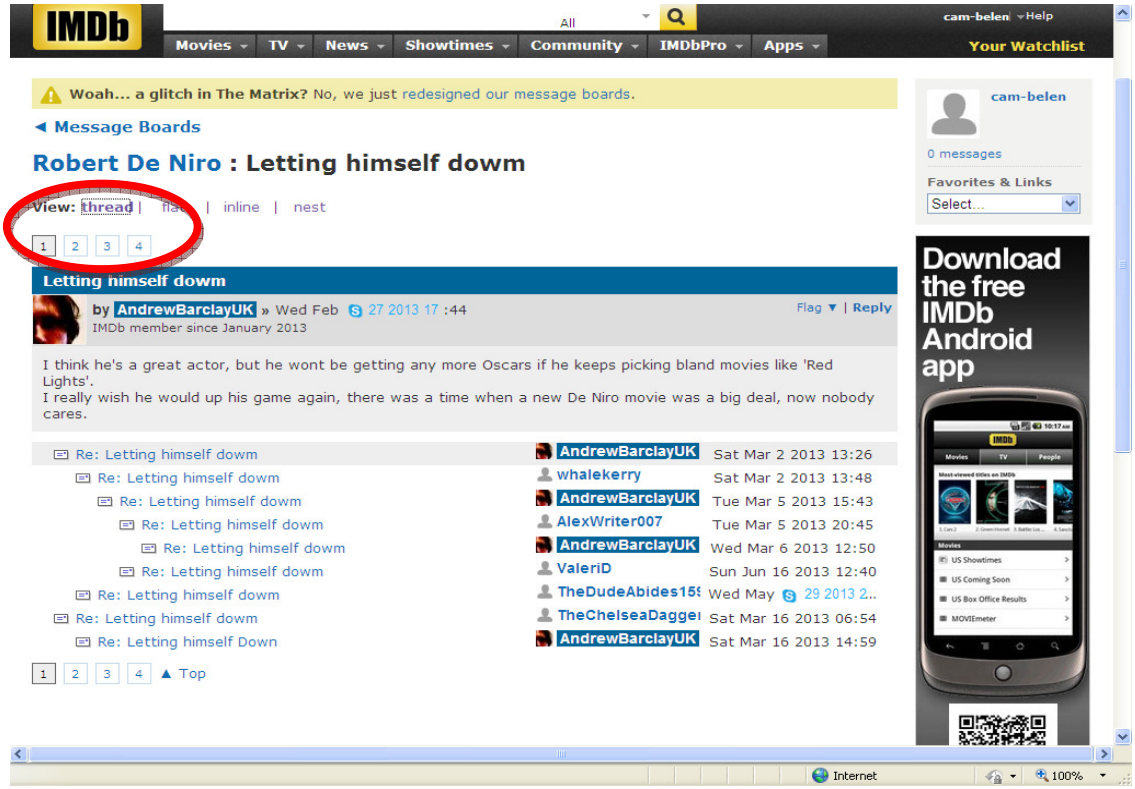


Figure 4.1: IMDb message board thread by *thread* view option

In Figure 4.1, it can be seen that the first post is given in full and the subsequent nine messages are represented by the subject line and date only. The second to the sixth posts all respond to the immediately previous post and can be considered to be the first sub-thread of the thread. The seventh post by ValeriD responds to the fourth post and can therefore be considered to be a second sub-thread. The eighth post by TheDudeAbides155 responds to the second post and so represents a third sub-thread. The second last post responds to the first post and the last post responds in turn to the second last post. The last two posts therefore represent a fourth sub-thread.

Figure 4.2: IMDb message board thread by nest view option

The *nest* view option was employed in the compilation of the IMDb corpus and the term *nest structure* is employed by this study to describe the turn structure of a thread (see Figures 4.3 and 8.1). However, as the thread data was converted to text-only format, the indentation information was lost but put back into the text-only file. As the corpus was to be annotated manually, this visual representation of the thread interaction structure would suffice. It was also necessary to reinsert certain aspects of the original message board text in the text specifically so-called *markup* functions, which add certain features to plain text. *IMDb* message boards have a number of these markup functions including a variety of icons, such as:

 [sad]
 [2face]
 [evil4]

 [fight2]

Most importantly for this study, it also includes the [quote] [/quote] markup function to indicate that a part of a previous poster's text will appear as a quotation in the present post. All markup functions were represented in the text-only file by the markup instruction in square brackets.

4.2.2 Size of the IMDb corpus and selection criteria

Given the intensive nature of the system of analysis proposed (as detailed in this chapter), it was decided to compile a small corpus of around 15,000 Linear Unit Grammar (LUG) elements to be manually annotated employing the *UAM CorpusTool* following an adaptation of the model of LUG described in section 2.3. An initial estimate of three words per element based on those elements provided by Sinclair and Mauranen gave a target of 45,000 words of text required. This figure compares favourably with similar manually annotated corpus-based studies employing the *UAM CorpusTool*. For instance, Martín-Úriz and Murcia-Bielsa (2008) annotated a corpus of less than 2,000 words for generic features in learners' writing; Lozano (2009) annotated a corpus of around 25,000 words of errors in Spanish learners of English; Moore (2011) annotated a corpus of around 10,000 words for discourse flow; and Thompson (2012) annotated a corpus of newspaper editorials of around 20,000 words for subjectivity.

As regards the selection of the threads, given the objectives of the thesis described in Chapter 1, any of the threads present on the *IMDb* website in theory would have been as

appropriate as any other. It was therefore decided to adopt a quasi-random approach to the compilation keeping in mind certain guiding principles and constraints. The first of these was that the most recent message in all threads should be within a year of compilation in order to guarantee some degree of currency. As the corpus was compiled in December 2006, all threads contained at least one post from the year 2006.[17] Secondly, the IMDb user charts were consulted as a means of gauging interest in films, directors and actors within IMDb; as was information from outside IMDb regarding films, actors and directors current at the time of compilation. The details regarding the threads that make up the IMDb corpus can be seen in Appendix 4.1. Each message board provided between one and four threads depending on their length.

4.2.3 Ethical considerations and online message board corpus compilation

The use of data from the Internet for analysis raises certain issues, particularly, as was the case here, where the data is used without the possibility of gaining permission from the participants themselves. Generally, according to the *Association of Internet Researchers* if the original text from the Internet is available to the general public then it is in the public domain and can be readily used in research (Markham and Buchanan 2012). This viewpoint is shared by previous researchers of message board discourse (Collot and Belmore 1996; Claridge 2007; Tanskanen 2007; Arendholz 2013) and by the present writer. Following Arendholz (2013), all usernames and real names were removed from the IMDb corpus data to protect anonymity.

[17] In the event, 787 posts in the IMDb corpus actually came from 2006. Two of the remaining were from 2005; one from 2004.

Having looked at the gathering and compilation of the corpus for this study, we can turn to the theoretical considerations which informed the development of the system of analysis.

4.3.1 The system of analysis

4.3.1 Introduction

There are two basic tenets of the system of the analysis in this study based on what has been reviewed in Chapters 2 and 3. First, following Sinclair (1981/2004c; 1985/2004d) discourse can best be described as a dynamic process in which the reader/listener is continually processing incoming language in a linear fashion. The second tenet is that discourse reflexivity can best be described as a quality that certain parts of the text possess whereby they make an explicit reference to the present discourse.

As a consequence of adopting a dynamic model in this study, the researcher takes on the role of a reader of the discourse and thus analyzes each unit of analysis in the discourse as the text of the moment (Sinclair 1993/2004e:82). In grammar, the text of the moment is the *element*, in discourse it is the *linear unit*. The text of the moment is conceived as being the focus of attention for the reader at that moment and changes from one unit of analysis to the next in a linear fashion. Fundamentally, the researcher analyzes the present unit for analysis as if s/he has read everything in the discourse up to that point but has not read beyond it. The researcher then tries to make a connection, if any, with the previous unit and decides if the present text of the moment prospects anything in the upcoming unit. Thus, this *in situ* perspective of the researcher attempts to recreate the

dynamism of the interaction of the reader as s/he proceeds through the text. This inevitably makes for a linear analysis as the researcher/reader is limited to relating the text of the moment to its immediate surroundings. What becomes paramount in such a model are considerations such as what will appear in linear sequence in the text immediately after the text of the moment, whether this is signalled or predicted in any way and how the text of the moment relates to both the text immediately preceding it and to the shared knowledge established between the participants up to this point. Thus, the notions of encapsulation, prospection and completion (see Chapter 2), are central to the model employed here. For that reason, adaptations of LUG and of the analytic system for written discourse structure (see section 2.5) were employed as the systems of analysis in this study. In addition, certain features of the outline model of spoken discourse structure (see section 2.4) were included. In particular, what was previously termed *challenge* in previous models of spoken discourse (Burton 1980; Sinclair 1992; Warren 2006), termed here *suspension*, takes on central importance in the system of analysis of this study.

In this way, an adaptation of LUG provides this study with a means of dividing the data into the smaller of two analytical units: the *element*, as well as a means of categorizing these elements. It also provides a principled means of delimiting the analytical unit used in the application of the models of spoken and written discourse. This unit is called the *linear unit.*[18] Thus, in this study, the LUG element categories of message-oriented (m) elements and two types of organizational (o) elements: interactive organizational (oi) and textual organizational (ot) elements and their subcategories are employed, with some

[18] The term *linear unit* will be used in this study. There are two categories of linear unit: M linear units (meaning linear units) and OI linear units (interactive organizational linear units).

modifications from LUG. In the analysis of linear unit relations, there are two main categories: M linear units and OI linear units as well subcategories of each of these.[19]

All of these are outlined and discussed in sections 4.3.2 and 4.3.3.

4.3.2 System of analysis for element relations

4.3.2.1 Introduction

This section will describe in detail the system of analysis for relations between the smaller of the two analytical units employed in this study: the element.

As described in section 4.3.1, two different conceptual frameworks were employed in the classification of elements and element relations. Firstly, an adaptation of the system of analysis used in LUG (as described in section 2.3) was employed. Secondly, having classified the elements employing this adapted LUG system of analysis, the elements were further classified as to whether they are discourse reflexive (DR) or non-discourse reflexive (non-DR).

Several additions and several deletions were made to Sinclair and Mauranen's categorization of LUG for this study. The first major addition is the further sub-classification of o elements. In LUG these remain simply OI and OT elements. In this study, oi elements have been further categorized as core and qualified oi elements and both qualified oi and ot elements have then been categorized into subcategories in a

[19] In this study lower case coding (m, ot , oi etc.) refers to the element whereas upper case coding (MS–, +M etc.) refers to the linear unit.

similar fashion to m elements. This will be presented in Steps 5 to 7. The second major addition to LUG is the further distinction made within qualified elements between compliant and suspensive elements. This will be presented variously in Steps 5 to 7. Deletions from LUG include the deletion of overlapping M elements, i.e. MA and MR elements. This will be discussed briefly in Step 4.

The seven steps undertaken for the categorization of the IMDb corpus data following the LUG-based system of analysis will now be described. In this chapter, examples are taken from Thread 017 from the IMDb corpus regarding the American actor Robert De Niro and whether posters feel that it is time that he should retire (hereafter known as the *De Niro* thread. See Appendix 4.2 for a complete analysis of element relations of the *De Niro* thread. [20]

4.3.2.2 Step 1: Dividing the text into elements[21]

The first stage of the analysis is the division of the text into elements based on the reader-based chunking of the text as detailed in section 2.3.

As Step 1 represents the most intuitive of all of the steps, it was decided to carry out a limited inter-reliability procedure for this study in order to bolster any subsequent claims made regarding the categorization of elements. In order to do this, a random extract of 288 words from Thread 043 was chosen. The extract was then divided into elements by two independent raters, myself and Dr. Martin Hewings, the initial supervisor of this

[20] The line references and participant letters for the *De Niro* thread follow those of Appendices 4.1 and 4.2. For other examples the extract starts at Line 1 and the participants at A.
[21] Space does not permit giving extensive examples of each step. However, the outcome of the steps is illustrated in examples 4.1 to 4.16. The outcome is also fully illustrated in Appendix 4.2.

thesis. The raters' proposed Provisional Unit Boundaries (PUBs) (see section 2.3.3) for the extract were then compared. This procedure produced a Kappa co-efficient of 0.92. Following the standard interpretation of the Kappa co-efficient provided by Landis and Koch (1977), a figure of 0.81 or above signifies 'almost perfect agreement' between raters. It was therefore concluded that the inter-rater reliability in this exercise for Step 1 was extremely high.

4.3.2.3 Step 2: Categorizing elements as m elements or o elements

The classification of the elements as message-oriented (m) elements action-oriented (o) elements follows Sinclair and Mauranen's definition, as described in section 2.3 and summarized as follows:

> **m elements**: message-oriented elements, which lead 'to the incrementation of shared experience' (Sinclair and Mauranen, 2006: 51);
>
> **o elements**: organizational elements, whose primary function is the management of the process of incrementation of knowledge through facilitating the organization of the text or facilitating the interaction.

A continuation of the inter-rater reliability procedure as described in section 4.3.2.2 for Step 1 was also carried out for Step 2 for the 93 elements established from the PUBs agreed upon in Step 1 for the same extract. In this instance, the same two raters independently categorized the elements as m or o elements. This produced a Kappa

coefficient figure of 0.89. This again can be interpreted as representing extremely high agreement between the raters.

4.3.2.4 Step 3: Categorizing m elements by status

The classification of the m elements based on their status – core or qualified – again follows Sinclair and Mauranen. A core status m element is one which is perceived as being a complete message or, in terms of syntagmatic relations, one which demonstrates no syntagmatic relation with the elements preceding or following it (see Chapter 2) within the same turn.[22] Qualified status m elements, on the other hand, are those where a syntagmatic relation between the m element and a preceding or upcoming element is evident.

4.3.2.5 Step 4: Sub-categorizing qualified-status m elements

Due to the fact that online message board discourse is both written and asynchronous, it was envisaged that the overlapping M elements – MR and MA elements – and the insubstantial MF elements (see Tables 2.5–2.7) would play little or no part in the IMDb corpus data, as they are, for the most part, related to real-time spoken interaction, where online adjustments, interruptions, disfluencies and so on occur. After the initial analysis of the IMDb data, no evidence was found of the presence of MA or MR-like elements, though occasionally elements which may be considered to be MF were found. It was therefore decided to maintain the MF element category as having a potential

[22] It should be noted that when considering element relations, this study limits itself to relations within the turn. Relations between turns are only considered when examining relations between linear units.

representation in the IMDb corpus discourse but to delete the categories of MA and MR elements.

The system of analysis therefore comprised the following categories of qualified status m elements[23]:

Compliant elements: m–, +m, +m–, ms, ms–

Suspensive elements: |m, |m–, |mf

Summaries of each of these can be seen in Table 4.1 and a discussion of those subcategories which have been added to LUG will follow.

Table 4.1 Subcategories of m elements

Type	**Symbol**	**Description**	**Examples[24]**
Core status elements	m	Complete message oriented element which either initiates a linear unit or re-initiates a linear unit following a suspensive element and which shows no relation with other elements	*he enjoys acting* (line 298)[25]; *Deniro still the man!* (line 256)
Compliant qualified status elements	m–	Incomplete message oriented element which either initiates a linear unit	***Pacino's last great performance*** **/** *was 2 years ago,* (**m– /** +m) (lines 191–192)

[23] In the coding of elements, a dash (–) indicates an incomplete element; a plus (+) indicates a completing element; and a bar (|) indicates that the element is suspensive.
[24] In cases where more than one element in a series is shown, the element in question is in bold
[25] In Tables 4.1, 4.2 and 4.3, examples with line references are from the *De Niro* thread. Examples from other threads in the IMDb corpus are indicated with the thread code.

		or re-initiates a linear unit following a suspensive element	
	+m	Message oriented element which completes incomplete element thus complying with prospection	*Pacino's last great performance* / ***was 2 years ago***, (m– / **+m**) (lines 191–192) *what does deniro* / *have to retire*/ *if he likes doing* / ***what he does***:[26] (m– / +m / ms– / **+m**) (lines 272–275)
	+m–	Message oriented element which partially completes incomplete element thus complying with prospection	*he should* / ***follow the example of*** / *good actors* (m– /**+m–** / +m) (lines 52–54)
	ms	Message oriented supplement to completed message element (i.e. m, +m or ms element)	*He has done* / *some pretty crap films* / ***in the last 5 years.*** (m– / +m / **ms**) (lines 22–24) *he wants to* / *try different roles*/ *and* / ***experiment*** (m– / +m / +ot– / **ms**) (lines 287–290)
	ms–	Incomplete message oriented supplement to completed message element (i.e. m, +m or ms element)	*what does deniro* / *have to retire*/ ***if he likes doing*** / *what he does:* (m– / +m / **ms–** / +m) (lines 272–275)
Suspensive qualified status elements	|m	A message oriented element which either initiates a linear unit in a way that does not comply with expectations or which follows an incomplete m element (i.e. m–, +m– or	***like a boxer*** / *even great actors* / *will work into their 80's* (|**m** / m–/ +m) (lines 31–33) *The most recent movies* / ***that DeNiro have been making*** / *had*

[26] Examples quoted from the IMDb corpus are reproduced exactly as were found including non-standard spelling and grammar.

		ms– element) in a way that does not comply with expectations	*been a joke!* (m– / **|m** /+m) (lines 201–203)
	|m–	Incomplete message oriented element which either initiates a linear unit in a way that does not comply with expectations or follows an incomplete m element (i.e. m–, +m–, ms– element) in a way that does not comply with expectations	**is it Time for him** / to Retire? (|**m–** / +m) (lines 18–19)
	|mf	Message fragment	*whether* / *there is* / *any fact to it* /***at*** / *or not* (ot / m– /+m / **mf** / ms) (005)

For the most part, then, LUG was followed for core and qualified compliant elements as illustrated by the examples provided in Table 4.1. As stated above, perhaps the most important addition in the present study's system of analysis is the addition of the distinction between compliant and suspensive qualified elements. Following Brazil (1995), this distinction is made in m elements based on whether the element complies or does not comply with the expected sequence of elements to reach the target state. Thus, a compliant message-oriented element fits at least one of the following criteria:

a. It initiates the linear unit as per the expected sequence, i.e. with a noun or noun-verb (NV) sequence (in the case of m– elements);

b. It completes an incomplete message-oriented elements as expected (in the case of +m and +m– elements);

c. It provides supplementary information to extend the completed message (in the case of ms and ms– elements).

Thus, in Example 4.1, we can see three compliant m elements in sequence from the *De Niro* thread. The first is compliant as it begins with a subject pronoun followed by a verb. It is incomplete as the element prospects a completion in the form of an object to the verb. The +m element in line 23 then completes that prospection. The ms element in line 24 provides supplementary information to a completed sequence of m elements in lines 22–23.

Example 4.1

P[27]	**L**	**Element**[28]	**Element type**
C	22	He has done	m–
	23	some pretty crap films	+m
	24	in the last 5 years.	ms

As explained in section 2.2.3, suspensive message-oriented elements can occur either in a linear unit-initial position or a linear unit-medial position. Those that occur in an initial position begin the linear unit in a way which breaks the expectation that a linear unit is

[27] P= Poster T= Turn and L= Line number. T only features when there are two or more turns.
[28] In element example tables, double horizontal line indicates the end of a turn; a solid horizontal line indicates the end of a subject line; a dotted line indicates the end of an element. A wavy line means that part of the same turn is omitted. A double wavy line means that part of the discourse is missing and that it is the end of a turn. The posts are divided into elements, one per line.

initiated with an NV sequence. Hence, a suspensive m element in initial linear unit position fits at least one of the following criteria:

d. It begins the linear unit with an open selector (*if*, *when* etc.);

e. It follows immediately after a linear unit initial lot element (*although*, *since* etc.);

f. It does not contain a finite verb (e.g. sentence adverbials);

g. It contains a noun or NV sequence but one that acts as a preface to the linear unit;

h. It contains a VN sequence rather than a NV sequence.

Suspensive elements in initial position, then, signal to the reader that the target state will not be reached by the end of the present element or the sequence of elements even though it reaches a point of apparent completion. This can be seen in Example 4.2, which represents category d. above, in that the element in line 1 begins with the open selector *if*. Hence in Example 4.2, when the reader reaches the end of line 1, the element is complete and yet the target state has not been reached. The target state has similarly not been reached by the end of line 2. Instead, the reader knows that it will take a second initiation within the same linear unit (hereafter called a *re-initiation*), starting in line 3 and completed in line 4, before the target state is reached.

Example 4.2[29] [30]

P	L	Element	Element type
A	1	If Audrey had died	|m
	2	in the 60's,	ms
	3	she would receive	m–
	4	the same kind of attention.	+m

Category e. is similar to Category d. in that the linear unit begins with an open selector. However, in this case the open selector is a separate suspensive |ot element. This can be seen in example 4.3.

Example 4.3

P	L	Element	Element type
A	1	although	|ot
	2	Audrey had her hardship	|m
	3	her life	m–
	4	was boring	+m
	5	in comparison	ots

Example 4.4 represents an example of Category f., i.e. a linear unit-initial |m element which does not contain any sort of verb, finite or otherwise. When the reader reaches the end of line 31, as we have seen in Example 4.2, some sort of a premise has been established. Nevertheless, the reader knows that a re-initiation is still necessary to reach the target state. The re-initiation begins in line 32.

[29] Examples 4.2 and 4.3 are taken from Thread 002 about the actresses Audrey Hepburn and Marilyn Monroe.
[30] In example tables, the element / linear unit which is referred to is shaded. DR elements are in bold.

Example 4.4

P	L	Element	Element type
E	31	like a boxer	lm
	32	even great actors	m–
	33	will work into their 80's	+m

Category g. above refers to those suspensive elements that follow an NV sequence but which signal to the reader that the target sequence will not be reached with the present element or element sequence and that a re-initiation will begin after the end of the element sequence. An example of this type can be seen in Example 4.5.

Example 4.5

P	L	Element	Element type
F	51	i think	lm
	52	he should	m–
	53	follow the example of	+m–
	54	good actors	+m

Such suspensive m elements often contain the personal pronouns *I*, *you* or *it*. They are also often followed by the +ot– element *that*, which, in this role, can be considered to be a signal that a re-initiation of the linear unit is about to begin. In other grammatical descriptions, these suspensive elements have been termed: a verb/adjectival predicate which controls a *that* clause (Biber et al 1999); a verb + clause complement; a V *that* and ADJ *that* pattern (Francis et al 1996, Francis et al 1998; Hunston and Francis 2000); a

reporting verb in indirect speech (Swan 2005); and an evaluative *that* construction, i.e. matrix clause + *that*-clause (Hyland and Tse 2005).

The final category of linear unit-initial suspensive elements, Category h., represents those which contain a VN sequence, i.e. those which have an interrogative form. This can be seen in Example 4.6a.

Example 4.6a

P	L	Element	Element type
A	18	is it Time for him	m–
	19	to Retire?	+m

In the present study, these elements are combined in the analysis of linear unit relations to form an M– linear unit (see section 4.3.5) with inter-turn orientation, i.e. an elicitation. A response from another participant then acts as the linear unit which provides the completion to the prospection. This can be seen in Example 4.6b.

Example 4.6b

P	T	L	Linear Unit[31]	Linear Unit type
A	1	4	is it Time for him / to Retire?	M–
B	2	6	NO!	+M

[31] In linear unit example tables, double horizontal line indicates the end of a turn; a solid horizontal line indicates the end of a subject line; a dotted line indicates the end of a linear unit. A wavy line means that part of the same turn is omitted. A double wavy line means that part of discourse is missing from another turn. The posts are divided into linear units, one per line.

As Brazil (1995) explains, the target state of an asking exchange is not achieved until a response is provided. Therefore, just as has been seen in Examples 4.2 and 4.3, a suspensive m element, as seen in Example 4.6a, is a signal to the reader that the target state will not be reached even when completion is reached at the end of the element in line 19.

Suspensive message-oriented elements in linear unit-medial position, on the other hand, are those elements which initiate a new unprospected sequence in the middle of a linear unit when they were expected to simply complete an incomplete element. An example of a suspensive m element in medial position can be seen in Example 4.7, where the initial element in line 201 prospects a completion through some sort of verb phrase. Instead, in line 202 a suspensive |m element appears, which suspends that prospection by providing supplementary information. However, this suspension is temporary and the prospection which is unfulfilled from line 201 is completed by the element in line 203.

Example 4.7

P	L	Element	Element type
A	201	The most recent movies	m–
	202	that DeNiro have been making	|m
	203	had been a joke!	+m

4.3.2.6 Step 5: Categorizing o elements by focus

Following Sinclair and Mauranen, o elements are divided into two main categories:

oi elements: those elements whose function is the management of the interaction;

ot elements: those elements whose function it is to manage the internal coherence of the text.

Examples from the *De Niro* thread of oi and ot elements can be seen in Table 4.2.

Table 4.2 o elements from the *De Niro* thread

oi elements		ot elements	
absolutely	Just tell me	and	However
anyone?	LOL	As of now	in much the same way as
anyway	man,	because	like
buddy	maybe	but	Now
C'mon	no	but in reality	or
dammit	of course	'Cause	Second
eh	ok	either	so
etc	people	even though	that
euh...	perhaps	First off	then
FACT	Really?		
Hands down!	ROFL		
hey	ROTFLMAO!!!		
hmm	So what if		
i got one...	To answer your question		
i mean	wait		
I wanna know	yeah		

4.3.2.7 Step 6: Categorizing oi elements by status

Unlike ot elements, certain oi elements can be considered to be core oi elements. These are oi elements which, on the one hand, are not primarily message-oriented in that they are better considered to be interactional in their orientation. However, they are also separate from the surrounding elements enough to be considered to have no syntagmatic

connection with those elements. In practice, these elements tend to be those which are stand-alone interactional responses to a previous turn as can be seen in Example 4.8.

Example 4.8[32]

P	T	L	Element	Element type
A	1	1	Thank you so much	oi
		2	for my laugh of the day!!!!	ms
		3		ois
B	2	4	You're welcome	oi
		5	:)	ois

In this example, we can see two core oi elements functioning in a rather formulaic phatic exchange. The first core oi element is seen in line 1 thanking Poster B for a previous post. This is clearly an interactional element rather than an information-oriented element. Similarly the second core oi element is a follow-up to the previous post, functioning as an F move does in the IRF exchange structure (see section 2.4). Again, this is clearly interactional rather than message-oriented.

Table 4.3: Core and qualified oi elements

Type	Symbol	Description	Examples
Core	oi	A complete interactional organizational element with no relation with surrounding elements within the same turn	*Hi,* (022) *Peace on earth* (013) *You're welcome* (013)

[32] This example is taken from Thread 013 about American actor Kevin Spacey. Poster B is thanking poster A for a previous post where Poster A ridicules the original poster for asking about Kevin Spacey's sexuality.

Compliant qualified status elements	ois	An interactional organizational element which provides supplementary interactive orientation regarding the previous m element sequence and an end to the sequence	*i couldnt agree with you more /* ***buddy***. (m/ ois) (lines 39–40)
Suspensive qualified status elements	loi	An interactional organizational element which either: a. initiates a linear unit and provides interactive orientation regarding the present linear unit	***perhaps*** */ you should / go watch Hide And Seek /again* (**loi** / m– / +m / ms) (lines 149–152) ***i mean,/*** *he made those two films* (**oi /** m) (lines 44–45)
		b. signals the suspension of the prospection created by an incomplete m element (i.e. m–, +m– or ms– element) in the middle of a linear unit and also provides interactive orientation regarding the present linear unit	*It has /* ***like /*** *finding Nemo / on it* (m– / **loi** /+m– / +m) (070)

In contrast to core oi elements, a qualified suspensive loi element in linear unit-initial position signals that the initiation of the linear unit has been suspended. At the same time, it provides a 'preliminary intimation' (Brazil 1995:184) as to the character of the upcoming linear unit and how it should be approached by the reader. This is done, for example, by acting as a hedge or booster (Hyland 2005) in a prelude to the linear unit as can be seen in Example 4.9, where the loi element *perhaps* acts as a hedging device, creating a 'cushioning effect' (Sinclair and Mauranen 2006:74).

Example 4.9

P	L	Element	Element type
E	149	perhaps	|oi
	150	you should	m–
	151	go watch Hide And Seek	+m
	152	again	ms

However, there are occasions where |oi elements in linear unit-initial position both provide a preliminary intimation about the upcoming element sequence and also signal a relation between the previous linear unit and the present linear unit. Thus, in Example 4.10, the |oi element *i mean* is employed to overtly indicate to the reader that there is a connection between the previous linear units (lines 41–43) with the upcoming linear unit (lines 44–45). In this case, it is a signal to the reader that the upcoming linear unit will be some sort of clarification.

Example 4.10

P	L	Element	Element type
A	41	this guy	m–
	42	hasn't made a good film	+m
	43	in along time.	ms
	44	**i mean,**	**|oi**
	45	he made those two films	m

Similarly, in a turn-initial position in Example 4.11, the |oi element *yeah* both indicates that the poster has understood that a response is required for the initial question but also

communicates that the upcoming element sequence in line 39 will be providing this response. As Stubbs (1981:110) comments, these elements are 'Janus-faced' in that they have both prospective and retrospective orientation (Aijmer (2002:37).

Example 4.11[33]

P	L	Element	Element type
A	1	Is It Time	|m–
	2	to Retire	+m
	3	already?	ms
F	38	yeah	|oi
	39	i couldnt agree with you more	m

Brazil (1995) asserts that such interactive elements do not aid in the sequence reaching its target state and as such are not prospected. In this study, this means that they are categorized as suspensive |oi elements. This can be seen most clearly in |oi elements in a medial position in the linear unit. In such a position the |oi element suspends the prospection made by an incomplete m element (m– element etc.). The |oi element interrupts the progress towards the target state by providing some sort of interactive orientation for the reader. In Example 4.12, the |oi *like* suspends the prospection created in the previous element in line 1 temporarily and provides an interactive interlude which represents an expression of vagueness. The prospection is completed in line 3.

Example 4.12[34]

[33] The turn which begins in Line 38 is in response to the initial Elicitation in lines 1–3.

P	L	Element	Element type
A	1	It has	m–
	2	like	loi
	3	finding Nemo	+m
	4	on it	ms

Nonetheless, there is one type of oi element (ois element), which is categorized as being compliant. These elements provide supplementary interactional orientation at the end of a linear unit as well as reiterating its completion. The ois element in Example 4.13, for instance, adds an interpersonal sense of collegiality to the interaction through the use of the vocative *buddy*.

Example 4.13

P	L	Element	Element type
F	39	i couldnt agree with you more	m
	40	buddy.	ois

4.3.2.8 Step 7: Categorizing qualified-status ot elements

Another addition to LUG in this study is the distinction made between different ot element types. By definition, all ot elements are qualified-status elements as they signal some sort of relation between the previous and upcoming elements. However, within qualified ot elements three subcategories are proposed here (as seen in Table 4.4).

[34] This example is from Thread 070 in which the posters question the high rating of the film *Fight Club* in the IMDb rating list. This poster is questioning the validity of the rating list by commenting that the children's animated film *Finding Nemo* features on the list.

Table 4.4: Subcategories of ot elements

Type	Symbol	Description	Examples
Compliant qualified status elements	+ot–	A text focused organizational element which signals a relation between a preceding m element and a following m element	*godsend **/and /** hide and seek.* (m/ **+ot–** / ms) (lines 46–48) *So what if / his career / lately/ hasn't been /as great as before, / **but** / it happens* (|oi /m– / |m / +m– /+m / **+ot–** / m) (lines 277–282)
	ots	A text focused organizational element which signals supplementary textual information about the previous m element sequence and the end to the sequence	*well / I liked the movie / **as well*** (|oi / m / **ots**) (065)
Suspensive qualified status elements	|ot	A text focused organizational element which either: a. initiates a linear unit and signals that the following m element sequence will not comply with expectations b. signals the suspension of the prospection created by an incomplete m element (i.e. m–, +m– or ms– element) in the middle of a linear unit	***although** / Audrey had her hardship / her life / was boring / in comparison* (**|ot** / |m / m– / +m / ots) (002) *he was not winters / **or** / anybody elses/ best friend* (m– / **|ot**/ |m– / +m) (064)

A +ot– element, therefore, is an organizational elements that shows a link between the previous and the upcoming element. This can be seen, for example in the ot element *but* in line 282 in Example 4.14. In this example the +ot– element signals a link between linear units, lines 277–281 being one linear unit and line 283 being the other.

Example 4.14

P	L	Element	Element type
O	277	So what if	|oi
	278	his career	m–
	279	lately	|m
	280	hasn't been	+m–
	281	as great as before,	+m
	282	but	+ot–
	283	it happens	m

The second type of compliant ot element is an ots elements, which provides supplementary textual information as well as signalling the end of the linear unit. These include *too*, *either*, *as well* and so on (see Example 4.15).

Example 4.15[35]

P	L	Element	Element type
A	1	well	|oi
	2	i liked the movie	m
	3	as well	ots

[35] This is from Thread 065 about *Battlefield Earth*, which heads the IMDb poll as the worst film of all time.

Just as we have seen that qualified status m elements can be divided into compliant and suspensive elements, the same is true for ot elements. Suspensive |ot elements either occur:

a. Initially in a linear unit. These then indicate that the expected linear structure is temporarily suspended and that the upcoming m element sequence is, in fact, not the initiation of the linear unit for real, but rather a prelude. The re-initiation of the linear unit occurs after the signalled m element sequence. Suspensive ot elements, in a linear-unit initial position are typically ot elements such as *because* or *although*. Example 4.3 from above has been repeated here for ease of reference.

Example 4.3 (*Repeated*)

P	L	Element	Element type
A	1	although	|ot
	2	Audrey had her hardship	|m
	3	her life	m–
	4	was boring	+m
	5	in comparison	ots

b. In a medial position in the linear unit. In this case, the |ot element suspends the prospection made by an incomplete m element. In Example 4.16, the incomplete m– element in line 1 prospects the completion of the noun phrase. Instead of a completion, in line 2, the prospection is temporarily suspended by the |ot element *or*. It also signals the inclusion of unprospected suspensive m element, seen in line 3. The completion in

line 4 may be considered to be either the completion of the incomplete element in line 3 or the suspended prospection in line 1.

Example 4.16[36]

P	L	Element	Element type
A	1	he was not winters	m–
	2	or	lot
	3	anybody elses	lm–
	4	best friend	+m

In the next section, how these elements are combined together to form linear units will be explained.

4.3.3 System of analysis for linear unit relations

4.3.3.1 Introduction

This section will describe in detail the system of analysis for describing relations between linear units. A categorization based on the same as that seen for element relations in section 4.3.2 is employed for this purpose. Thus, as will be seen below, the concepts of prospection and completion play a central role in the system of analysis for linear unit relations just as they did in the system for element relations. These are supplemented with encapsulation, overlay and verbal echo from the analytic system for written discourse structure (Sinclair 1993/2004e). However, several important modifications are made to the original model proposed by Sinclair. Firstly, the analytical unit employed in this study

[36] This example is from Thread 064 about the TV mini series *Band of Brothers* in which posters discuss the character Herbert Sobel and his relationship to the other characters.

is the *linear unit* rather than the sentence (or s/m). This allows the analysis of the element to define the analytical unit employed in the analysis of discourse. Also, the same basic codification will be employed (albeit in upper case). Hence the categories of M and OI linear units and their subcategories are employed. Finally, the fact that IMDb corpus discourse is dialogic/polylogic discourse was taken into account by incorporating the concept of *turn orientation* into the description of each linear unit (i.e. whether the linear unit is linked to another linear unit within the present turn or to a linear unit in another turn). The dialogic nature of the IMDb corpus discourse also means that the concept of *suspension* seen in the analysis of element relations plays an important role.

In a similar fashion to section 4.3.2, the six steps undertaken for the categorization of the data following the analysis for the linear unit orientation will be outlined (see Appendix 4.3 for a full analysis of the *De Niro* thread).

4.3.3.2 Step 1: Combining elements into linear units

Following Sinclair and Mauranen (2006), elements are combined together to create linear units. There are two basic types of linear units: those with an m element as the core and those with an oi element as the core. The guiding principles which are adhered to when combining elements to form these two types of linear units are as follows:

a. Linear units with an m element as core

1. It must contain either an m core element, or an m– / (+m–) / + m combination, known as a cleft-core element sequence (as seen in

Example 4.17). A linear unit can only contain one core status m element or cleft-core element sequence.

Example 4.17

P	L	Linear unit	Structure (elements)
F	12	this guy / hasn't made a good film / in along time.	m–/+m /ms

As can be seen in this example, a linear unit with an m element as core may also contain related supplementary qualified status m elements (i.e. ms, ms– elements (with associated completion +m and +m– elements) and/or ois and ots elements.

2. In cases where ot and oi elements fall between the core and qualified status m elements, the o elements are incorporated into the linear unit (see Example 4.18).

Example 4.18

P	L	Linear unit	Structure (elements)
F	14	godsend / and / hide and seek.	m / +ot– / ms

3. In cases where suspensive m elements are used in linear unit-initial position, the sequence which is signalled by the suspensive m element is combined with the next sequence containing a core m element or cleft-core element sequence. This can be seen in

Example 4.19, where the |m element *like a boxer* is followed by a cleft-core sequence.

Example 4.19

P	L	Linear unit	Structure (elements)
E	9	like a boxer / even great actors / will work into their 80's / and / ridicule their legacy,	|m / m– / +m / +ot– / ms

4. In cases where suspensive m elements are used in linear unit-medial position, the sequence which is signalled by the suspensive m element is incorporated into the linear unit. This can be seen in Example 4.20, where the |m element *that DeNiro have been making* features between m– and +m elements in the linear unit.

Example 4.20

P	L	Linear unit	Structure (elements)
A	52	LOL / The most recent movies / that DeNiro have been making / had been a joke!	|oi / m– / |m/ +m

5. If an o element starts the linear unit, whether it be suspensive (|ot or |oi element) or compliant (+ot– element) it is also incorporated into the linear unit. This is seen in Examples 4.20, which begins with a |oi element and 4.21, which begins with a |ot element. [37]

Example 4.21

[37] This example is the linear unit equivalent of Example 4.3 already presented above.

P	L	Linear unit	Structure(elements)
A	1	although / Audrey had her hardship / her life/ was boring / in comparison	lot / lm / m– / +m/ ms / ots

b. Linear units with an oi element as core

1. It must contain an oi element as its core. It must contain one core oi element only, as can be seen in Example 4.22. These linear units often comprise one element, i.e. simply a core oi element.

Example 4.22

P	L	Linear unit	Structure (elements)
G	18	**ROTFLMAO!!!**	**oi**

2. An OI linear unit may also include accompanying supplementary ois and/or ms elements. In Example 4.23 for instance, the linear unit in line 1 is supplemented by an ms element and an ois element. The linear unit in line 2 is supplemented by ois element in the form of an emoticon.

Example 4.23[38]

P	T	L	Linear unit	Structure (elements)
A	1	1	Thank you so much / **for my laugh of the day!!!!** / 	oi / **ms** / ois
B	2	2	You're welcome / :)	oi / ois

[38] This is the same extract as seen in Example 4.8 above.

4.3.3.3 Step 2: Categorizing linear units as M or OI linear units

In the vast majority of cases, linear units with an m element as core are classified as M linear units and equally linear units with an oi element as core are classified as core OI linear units. However, occasionally a linear unit which has a core m or cleft core m sequence can be classified as being an OI linear unit, if the linear unit is acting more as an interactional unit rather than as a message incrementation unit. This can be seen in Example 4.24, in which the linear unit in line 2 is constructed as if it were an M linear unit with a core-cleft sequence. However, the function of the linear unit is primarily to hedge the opinion expressed in the previous linear unit.

Example 4.24[39]

P	L	Linear unit	Type	Structure (elements)
A	1	Because / he deserves another oscar…	MS	+ot– / m
	2	Of course / this is just / my opinion / :)	OIS	loi / m– / +m/ ois

4.3.3.4 Step 3: Categorizing M linear units by status

The same principles and means of categorization as that used in the analysis of element relations, seen in Section 4.3, are applied. Hence, M linear units are divided between core and qualified-status linear units.

Table 4.5: Types of M linear units

Type	Symbol	Description

[39] This example is taken from Thread 037 about American actor Tom Hanks in which whether he should win another Oscar is discussed.

Core status M linear unit	M	Complete linear unit of meaning which displays no structural link with previous or upcoming linear unit
Qualified status compliant M linear unit	M–	Incomplete linear unit of meaning which displays no structural link with previous linear unit
	+M	Linear unit of meaning which completes previous incomplete linear unit of meaning
	+M–	Incomplete linear unit of meaning which partially completes previous incomplete linear unit of meaning
	MS	Supplement to linear unit of meaning
	MS–	Incomplete supplement to linear unit of meaning
Qualified status suspensive M linear unit	|M	A complete linear unit of meaning which does not comply with linear expectations
	|M–	An incomplete linear unit of meaning which does not comply with linear expectations

A **core status M linear unit**, as can be seen in Table 4.5, is a complete linear unit which neither has an overt link with the preceding linear unit nor does it prospect a further linear unit. This can be seen in Example 4.25, where there is no overt linguistic link between the linear unit in line 16 and the previous linear unit in line 15, as would be the case if a +ot– element such as *so* or *therefore* were present. Sinclair (1993/2004e:96) argues that in cases where there is no linguistic structural link, coherence can be interpreted by the reader by asking what kind of relationship may be inferred from all of the information available. However, in this study, cases where no linguistic link is evident are classified as core linear units. This means that even though there may be some notional link between the linear unit and the preceding or subsequent linear unit, if the link is not linguistically signalled, it remains as a core M linear unit. Neither does the linear unit in line 16 overtly seek the contribution of another poster in order to complete the meaning.

Instead, it remains as a stand-alone conclusion to the argument presented in the rest of the turn. As will be seen in Chapters 6 and 8, these stand-alone M linear units at the end of turns may be regarded as a feature of IMDb corpus discourse.

Example 4.25

P	L	Linear unit	Type
F	13	**i mean,** he made those two films	MS–
	14	godsend and hide and seek.	+M–
	15	they were both the same goddamn movie.	+M
	16	i think he should follow the example of good actors like hugh grant and jude law and choose his roles a bit better	M

Qualified M linear units, in contrast, are those that display some sort of linguistic link between each other. These will be explained and illustrated in Step 4 below.

4.3.3.5 Step 4: Sub-categorizing M linear units by linear relations and turn orientation

Qualified status M linear units are divided into the following basic categories:

Compliant linear units: M–, +M, +M–, MS, MS–

Suspensive linear units: |M, |M–

Brief descriptions of each of these can be seen in Table 4.5. More detailed explanations of how these linear units manifest themselves in the linear structure of the discourse will be provided firstly for compliant linear units followed by suspensive linear units.

Compliant M linear units

Firstly, compliant M linear units which demonstrate relations within the same turn will be illustrated. These are referred to as having *intra-turn orientation*. Secondly, those which demonstrate relations between turns will be illustrated. These will be referred to as having *inter-turn orientation*.

Compliant M linear units with intra-turn orientation

An **M– linear unit** is an incomplete linear unit which displays no structural link with a previous linear unit. This type of linear unit, then, has two main properties. Firstly there is no overt linguistic link to the previous linear unit, and in this way it resembles those core M linear units identified in Step 3 above. Secondly, the poster makes some sort of commitment through a prospection that there will a completion provided in the next linear unit. It is therefore incomplete in itself. In Example 4.26, in terms of retrospective relations, the linear unit in line 42 can be said to be generally related to the topic of the previous linear unit but there is no overt linguistic link with the previous linear unit as such. It is therefore considered to be an M– linear unit. If we consider its prospective relations, the poster commits to providing an elaboration of the first linear unit in line 42, i.e. by describing a film or films that Pacino has made. The prospection is made here through the use of the DR element *Let's take a look at*. It is also made with the use of punctuation with the colon indicating that it is incomplete. It is a case of what has been termed *advance labelling* (Tadros 1985, 1994; Sinclair 1993/2004e) in previous studies. This linear unit can therefore be described as having a pro intra-turn orientation.

Example 4.26

P	L	Linear unit	Type
E	41	but perhaps you should go watch Hide And Seek again eh	+M
J	42	**Let's take a look at** what Pacino's done in the last 10 years:	**M -**
	43	Donnie Brasco, The Devil's Advocate, The Insider, Insomnia, Angels in America, People I Know, The Merchant of Venice, etc.	+M

A **+M linear unit** with intra-turn orientation is a linear unit which provides the completion to which the poster has committed in the previous incomplete linear unit. The completion in Example 4.26 is provided with a list of films. The linear unit in line 42 can therefore be categorized as having a pro intra-turn orientation as it refers forward to the upcoming linear unit within the same turn and the linear unit in line 43 can be categorized as having a retro intra-turn orientation as it refers back to the previous linear unit within the same turn.

A **+M– linear unit** displays a combination of the properties of the two linear unit types described above in that retrospectively it provides a completion and prospectively it makes a commitment to continue. This can be seen in line 2 of Example 4.27. Line 1 represents a suspensive response to a previous turn but it is a response that prospects some sort of elaboration. This is provided by the +M– linear unit in line 2. However, this linear unit in itself sets up a further prospection in that it states what is not the case and therefore would seem to prospect a statement in the subsequent linear unit of what the

case actually is. This is provided in line 3. This linear unit can therefore be described as having both a pro and retro intra-turn orientation.

Example 4.27[40]

P	L	Linear unit	Type
A	1	I have to disagree with your point #3	IM–
	2	It was NOT the rejection that made Col Fitts kill Lester	+M–
	3	It was the fact that he had revealed his tru nature in error.	+M

An **MS linear unit** provides a supplement to the previous linear unit. This linear unit characteristically provides supplementary information or an evaluation of the previous linear unit.

Example 4.28

P	L	Linear unit	Type
N	74	Deniro can play a diverse range of roles /and gives any movie an immediate sense of grandeur.. in much the same way as Al Pacino..	M
	75	keep up the awesome work!!	MS

As can be seen in Example 4.28, the MS linear unit in line 75 provides a retrospective evaluation of what is described in line 74. It can therefore be categorized as having a

[40] This example is taken from Thread F05 about the film *American Beauty*, where the homosexuality of one of the character is discussed.

retro intra-turn orientation as it refers back to the previous linear unit within the same turn.

An **MS– linear unit**, on the other hand, is a supplementary linear unit which is incomplete. An example of this type of linear unit with intra-turn orientation can be seen in line 2 of Example 4.29. Line 2 provides an additional criticism of the book. This is not prospected in line 1. The connection is provided clearly by the deictic encapsulation *Not only that*. However, the fact that it is a new point introduced would seem to prospect some sort of elaboration of this new point by providing more details as to how the book was badly written. This is provided in line 3.

Example 4.29[41]

P	L	Linear unit	Type
A	1	As we all know, it was all made up	MS
	2	Not only that, it was very poorly written	MS–
	3	Just judging by the few pages I looked at it's obvious that the author of this work of "fiction" never learned proper grammar, spelling, or punctuation	+M

Compliant M linear units with inter-turn orientation

Referring back to Table 4.5, qualified M linear units can also have inter-turn orientation. With the involvement of two discourse participants, the notion of linear expectations shifts. Whereas linear expectations in linear units with intra-turn orientation are used by

[41] This example is from Thread 005 about English actor Cary Grant. Here the poster is criticizing a biography of the actor.

both to facilitate communication and to achieve desired rhetorical effects, in sequences with inter-turn orientation, linear expectations can be conceived as what the respondent is expected to do to fall within the expectations of the particular exchange or as a means of controlling the discourse (Sinclair 1992). Those linear units with inter-turn orientation which are incomplete (M–, +M– and MS– linear units) can be categorized as having pro inter-turn orientation whilst those linear units which link retrospectively (i.e. +M, +M–, MS and MS– linear units) can be said to have retro inter-turn orientation.

An **M– linear unit** with pro inter-turn orientation, then, is a linear unit in which the poster provides a linear unit that requires completion from another poster to reach the target state. In non-technical terms, this often takes the form of the first part of a question-answer sequence, or in the terms of other related models, the first increment in an Asking Exchange (Brazil 1995) or the first move in an eliciting exchange (e.g. Sinclair and Coulthard 1975; Francis and Hunston 1992). An example of an M– linear unit with inter-turn orientation is provided in Example 4.30. The completion is provided by the **+M linear unit** in line 2. As these linear units have inter-turn orientation, this second linear unit is provided by another poster. The sequence is complete since the prospection created in line 1 is fulfilled by the response in line 2. This linear unit can therefore be said to have retro inter-turn orientation as it refers back to the previous linear unit in another turn.

Example 4.30[42]

[42] This example is from Thread 018 about American actor Robert De Niro.

P	T	L	Linear unit	Type
A	1	1	Robert's Freakiest role????	M–
B	2	2	I think it's gotta be hide and seek	+M

An **MS linear unit** with inter-turn orientation follows a linear unit from another poster which does not prospect a further linear unit to complete the sequence. In order for it to be considered to be an MS linear unit it should provide a supplementary statement related to the previous post. The discourse expectation is that, if a supplementary linear unit is provided by another poster, it is done so as to agree with, comply or support the previous poster's turn (see section 2.4.4). If this expectation is fulfilled it can be considered to be compliant, i.e. a supporting move in Burton's (1980) terms.

Example 4.31[43]

P	T	L	Linear unit	Type
A	1	1	I thought it was notthat bad	+M
		2	I think mostly because I loved the book so much	MS
B	2	3	I liked it too after watching it a couple of times	MS

Hence in Example 4.31, Poster B provides a similar sentiment in line 3 as is seen Poster A's statement in line 2. Such a supplementary statement showing agreement is categorized as being an inter-turn MS linear unit. The linear unit in line 3 can therefore be categorized as having a retro inter-turn orientation as it refers back to the previous linear unit in another turn.

[43] This is from Thread 065 about *Battlefield Earth*, which heads the IMDb poll as the worst film of all time.

The remaining two linear unit types (MS– and +M–) are 'Janus-faced' in that they represent both types of inter-turn orientation, i.e. pro and retro inter-turn orientation.

An **MS– linear unit** with retro inter-turn orientation is a linear unit which provides support for the previous poster's linear unit but itself also prospects some sort of completion. In Example 4.32 the MS– linear unit has retro inter-turn orientation but pro intra-turn orientation as it prospects a completion in the shape of some sort of contrast to the concession made in line 2 from the same poster. This is duly provided in line 3.

Example 4.32[44]

P	T	L	Linear unit	Type
A	1	1	That is exactly the sort of crap the director would spin to relieve themselves of any responsibility for the movie's consequences	MS
B	2	2	That might be true,	MS–
		3	apart from the fact that Fincher didn't write the story	+M

An example of a +M– linear unit which has the opposite orientation to the example in Example 4.33, i.e. retro intra-turn and pro inter-turn orientation, can be seen in Example 4.33. In this example, the linear unit in line 1 prospects some sort of contrast. This is signalled by the qualified positive evaluation seen in the use of the word *certain*. From this, the reader anticipates some sort of contrast. This is provided in line 2 and is signalled by the +ot– element *But*. However, the linear unit in line 2 is also an elicitation

[44] This example is from Thread 070 which is about the film *Fight Club*.

featuring in the first turn of the thread. It therefore has a pro inter-turn orientation and is eventually completed by several different responses to the question.

Example 4.33[45]

P	L	Linear unit	Type
A	1	He is excellent in doing certain characters.	M–
	2	But is it right **to call him** "the most versatile actor"?	**+M–**

Suspensive M linear units

In the previous part, compliant linear units with both intra and inter-turn orientation have been presented. These occur on occasions when linear expectations are adhered to. However, the question arises as to how the structure of discourse is affected when a poster decides to flout these expectations. One of the key claims of this study is that the same basic linear mechanisms can be perceived as being active in linear unit relations as in element relations. As has been seen in section 4.3.2, the concept of suspension has been applied to element relations in this study and will therefore be applied to linear unit relations.[46] There are two suspensive M linear unit categories: |M and |M– linear units. The first of these is complete whilst the second, the |M– linear unit, prospects some kind of completion. Both of these can have retro intra or inter-turn orientation. In the case of a |M– linear unit, it can also have pro intra or inter-turn orientation. Both linear unit types

[45] This example is from Thread 001 about the American actor Al Pacino.
[46] The term *suspension* is not entirely felicitous for linear unit relations in that it implies a temporariness to the cancellation of linear expectation when it may in fact turn out to be permanent. A distinction could have been made between *suspension* and *termination*. However this would have meant taking into consideration two relations and therefore would have meant a move to a more synoptic analysis.

will be illustrated below firstly with intra-turn orientation and secondly with inter-turn orientation.

Intra-turn orientation

A |M linear unit with intra-turn orientation undermines the previous linear unit either: by indicating that the previous linear unit has not reached the target communicative state satisfactorily; by contradicting the veracity of the previous linear unit; or by questioning the appropriateness of the previous linear unit.

Example 4.34[47]

P	**L**	**Linear unit**	**Type**
A	1	Petition for Dark Angel movie	M
	2	**This isn't really a petition,**	**|M–**
	3	**just reply once** if you love "Dark Angel"	**+M**

Example 4.34 begins with the subject line of the thread in line 1 in which the poster declares that the thread is to act as a petition for the TV series *Dark Angel* to be made into a film. However, in line 2 that assertion is contradicted in that the poster states that the thread is not a petition after all and is rather, a means for posters to show that they love the TV series. The contradiction is seen in terms of linear structure as a suspension in line 2 of the M Linear Unit in line 1. It therefore represents an example of a |M linear unit with retro intra-turn orientation. As can be imagined, such real-time correction and

[47] The example is from Thread 047 from a message board dedicated to the American director James Cameron.

reformulation is a common feature of face-to-face spoken interaction. However, it is unusual in asynchronous written discourse.

A slightly different type of |M linear unit with intra-turn orientation can be seen in Example 4.35. The linear unit in line 1 acts as a concession and therefore strongly prospects that a contrast will follow. It is therefore an M– linear unit. However, this initial prospection is not completed as expected. Instead it is left unfulfilled, at least temporarily and a |M linear unit intervenes in line 2 providing supplementary information rather than the prospected contrast. This suspension is signalled by the |ot *and.* The prospected contrast is eventually introduced in line 3 and signalled by the +ot– *but*. The linear unit in line 3 can therefore be seen as resuming the prospection which was temporarily suspended in line 2. The linear unit in line 2 can be described as having retro intra-turn orientation. This raises the question for how many linear units a prospection can remain active if a suspension intervenes. Sinclair and Mauranen (2006:83), when considering incomplete topic increments describe them as having 'a short life span' and that if the completion re-appears too late afterwards that it should be considered to be a new start. In this study, for ease of application, the maximum period that a prospection can remain active is if there is one intervening suspensive linear unit, as is seen in Example 4.35.

Example 4.35[48]

[48] This example is from Thread 058 about the comedy film *Borat*. The poster provides an opinion about a scene in which black squares are used to cover parts of the actors' naked bodies.

P	L	Linear unit	Type	Structure (elements)
A	1	It wasn't very pleasant	M–	m
	2	and / I'm glad / there was /liberal use of / the black squares	|M	|ot / |m / m– / +m– / +m
	3	but / it was hilarious.	+M	+ot– / m

A |M– linear unit with intra-turn orientation is a linear unit which similarly indicates that the previous linear unit has not reached the target communicative state satisfactorily but is itself incomplete. This can be seen in line 2 of Example 4.36, where the |M– linear unit challenges the veracity of his/her own previous linear unit in line 1. It is therefore suspensive. However, this is done in the form of an elicitation, though not an elicitation that the reader expects that another poster will reply to. Instead, the same poster provides the rest of his/her opinion in the form of a completion of the prospection made in line 2. This is prediction through a *question* in Tadros' terms (see Table 2.2).

Example 4.36[49]

P	L	Linear unit	Type
A	1	that's when u know it's unanimous.	M
	2	Why is it unanimous?	|M–
	3	simple. pacino is an average actor	+M

Inter-turn orientation

As can be seen from the two examples above, the motivation for suspension within the turn can often be characterized as the poster trying to achieve a certain rhetorical effect.

[49] This example is from Thread 001 about the American actor Al Pacino.

In contrast, suspensions between posts (i.e. in inter-turn orientation), by definition, are more related to posters dealing with the here-and-now of interaction. In essence, an inter-turn suspension is similar to what was termed *challenge* in previous models of spoken discourse (e.g. Burton 1980; Sinclair 1992; Warren 2006) as described in section 2.2.4. In inter-turn relations, then, |M and |M– linear units with inter-turn orientation indicate that the initial statement provided by the first poster is challenged in terms of either it not reaching the target communicative state satisfactorily; its veracity; its appropriateness or, in more extreme cases, the actual right the poster has to make such a statement.

An example of a **|M linear unit** with retro inter-turn orientation can be seen in Example 4.37, where Poster A in line 68 makes a statement reiterating his/her opinion at the end of a turn. Poster M then contradicts this statement in line 69. It therefore flouts the supposition that posters will support the previous statement and so is suspensive. It can be categorized as having a retro inter-turn orientation as it refers back to the previous linear unit in another turn. This type of inter-turn suspension will be discussed extensively in Chapter 8.

Example 4.37

P	T	L	Linear unit	Type
A	22	68	he needs to either retire or need to pick roles his age, dammit.	+M
M	23	69	Deniro still the man!	|M–
		70	he can pull out any roll at his age	+M

A similar example can be seen in Example 4.38. However, the difference between this example and that seen in Example 4.37 is that it contains an inter-turn exchange with an intervening repetition in line 25 in the form of a quotation repeated from line 9 of the thread. In this example, Poster I clearly contradicts the opinion expressed in line 25 in line 26 and so represents a suspensive |M– linear unit.

Example 4.38

P	T	L	Linear unit	Type
E	5	9	like a boxer even great actors will work into their 80's and / ridicule their legacy,	+M
I	9	25	"even great actors will work into their 80's and ridicule their legacy"[50]	|OI
		26	**First off** their legacy is not tarnished	**|M–**

An example of a **|M– linear unit** with retro inter-turn and pro intra-turn orientation can be seen in Example 4.39. In this example, the linear unit in line 76 follows an M– linear unit with pro inter-turn orientation in line 1. The elicitation is therefore suspended since Poster O does not provide a compliant response. In doing so, Poster O goes beyond simply answering the elicitation in line 1 with a compliant +M linear unit such as *No*. Instead, s/he questions the worth of the original question by refusing to provide a compliant response. S/he responds with what can be called, in non-technical terms, a rhetorical question. This linear unit can therefore be categorized as having a retro inter-turn orientation as it refers back to the previous linear unit in another turn. The linear unit

[50] The poster did not use the [quote] [/quote] markup code to create this quotation. Instead s/he simply used quotation marks. The categorization as a suspensive M linear unit containing a verbal echo is the same.

in line 76 is incomplete as it prospects the specification of the general phrase *he likes doing / what he does*. The use of the colon at the end of the linear unit also indicates that it is incomplete, just as was the case in line 42 in Example 4.26. The completion is provided in line 77 by the one-word linear unit *acting*.

Example 4.39

P	**T**	**L**	**Linear unit**	**Type**
A	1	1	Is It Time to Retire Already?	M–
O	25	76	what does deniro have to retire if he likes doing what he does:	M–
		77.	acting	+M

Hence, the inclusion of information about turn orientation means that this study provides a considerable amount of extra information about each linear unit over and above the linear unit labels seen in Table 4.5. Each linear unit is thus further classified according to both its retro and pro turn orientation. This can be summarized in Table 4.6, where a tick (√) indicates that the M linear unit subcategory has potentially that particular type of turn orientation and a cross (X) indicates it does not.

Table 4.6: M linear units by status and turn orientation

Type	**Symbol**	**Turn orientation**			
		Pro		**Retro**	
		Intra-turn	Inter-turn	Intra-turn	Inter-turn
Core M linear unit	M	X	X	X	X
Compliant M	M–	√	√	X	X

linear unit	+M	X	X	√	√
	+M–	√	√	√	√
	MS	X	X	√	√
	MS–	√	√	√	√
Suspensive M linear unit	|M	X	X	√	√
	|M–	√	√	√	√

4.3.3.6 Step 5: Sub-categorizing OI linear units by linear relations and turn orientation

From the evidence of the IMDb discourse, it would seem that all OI linear units can be classified as being qualified. In comparison to the number of subcategories of M linear units, the number of OI linear units is relatively fewer with only two categories, as seen in Table 4.7.[51] [52] Firstly, compliant OI linear units will be illustrated followed by suspensive OI linear units.

Table 4.7: Types of OI linear units

Type	Symbol	Description
Qualified status compliant OI linear unit	OIS	Interactional supplement to linear unit of meaning
Qualified status suspensive OI linear unit	|OI	A linear unit of interaction which does not comply with linear expectations

[51] It is possible to find examples of OI core linear units in other text-types For example a T-shirt simply with the word *AH! (http://www.badideatshirts.com/?gclid=CPzNp_31mrgCFc9g4godYxIAjA)*

[52] Phatic exchanges entirely made up of OI linear units are common in face-to-face spoken interaction. For instance a *greeting exchange* (Francis and Hunston 1992), such as :Speaker A: *Hi!* Speaker B: *Hi!,* could be classified as OI- / +OI since a prospection /completion relationship is evident. However, in IMDb corpus discourse greetings do not seem to prospect a phatic response, rather they act as preludes for their own turn. They are therefore categorized as |OI linear units with pro intra-turn orientation.

In intra-turn orientation, **OIS linear units** provide interactional support for the previous statement. For instance in Example 4.40, the OIS linear unit *FACT* in an M / OIS sequence reiterates the veracity of the statement.

Example 4.40

P	L	Linear unit	Type
Q	88	De Niro is the most admired actor currently working today.	+M
	89	FACT.	OIS

An **OIS linear unit** with retro inter-turn orientation, on the other hand, provides support in the form of a brief interactional acknowledgement for what was stated by the previous poster. This can be seen in Example 4.41. These can be considered to be similar to an *endorse* act at the head of an *acknowledging* move in Francis and Hunston (1992).

Example 4.41[53]

P	T	L	Linear unit	Type
A	1	1	I'm gay (and Australian. not that it matters) and loved the film.	M
B	2	2	**Well said** <name>	**OIS**

A suspensive **|OI linear unit with intra-turn orientation** can fulfil different roles. In the example in Example 4.42 the |OI linear unit in line 4 follows an M– linear unit in line 3. The |OI linear unit signals the temporary suspension of the prospection in line 3. The

[53] This example is from Thread 057 about the film *Borat* and its supposed homoerotic content.

elicitation therefore remains unanswered momentarily as the poster follows up the elicitation in line 4 with an interactive encouragement to other potential posters to contribute.

Example 4.42

P	L	Linear unit	Type
A	3	**Just tell me, people,** is it Time for him to Retire?	**M–**
	4	I wanna know.	|OI

In Example 4.43, the |OI linear unit with retro intra-turn orientation follows an M linear unit. Here the |OI linear unit has the opposite effect to the OIS linear unit in Example 4.40. Rather than bolstering the previous linear unit, the |OI linear unit in line 4 undermines the veracity of the statement. In this case, this is done for comic effect.

Example 4.43[54]

P	T	L	Linear unit	Type
A	1	1	Am I the only teenage boy who finds her unattractive?	M–
B	2	2	No,	+M–
		3	there are quite a few gay teenagers.	+M
		4	**Just kidding**	**|OI**

A **|OI linear unit with inter-turn orientation** is one in which the second participant protests or challenges the previous statement with a brief interactional acknowledgement.

[54] This example is taken from Thread 025 about American actress Angeline Jolie.

In Example 4.44, the |OI linear unit follows an M linear unit. The second poster uses an offensive |OI linear unit to protest against the previous statement. These can therefore be considered to be similar to a *protest* act at the head of an *acknowledging* move in Francis and Hunston (1992).

Example 4.44[55]

P	T	L	Linear unit	Type
A	1	1	You're a noob and not special for noticing it	M
B	2	2	Hey, *beep* you peace of *beep*	|OI

A |OI linear unit with retro inter-turn orientation can also follow an M– linear unit in a previous turn. In Example 4.45, this is done using an OI linear unit containing an imitation of a paralinguistic action **Sigh**....By refusing even to answer the question using an M linear unit, Poster B successfully expresses disdain for the original question and Poster A.

Example 4.45[56]

P	T	L	Linear unit	Type
A	1	1	Is it true he was bisexual?	M
B	2	2	***Sigh*…**	**|OI**

[55] This example is taken from Thread 069 about the film *Fight Club* in which the original poster claims to have spotted a small detail in the film.
[56] This example is from Thread 005 about English actor Cary Grant.

4.3.3.7 Step 6: Sub-categorizing retro linear units by retrospective supplementary mechanisms

Following Sinclair (1993/2004e), each relation between linear units described above can be further categorized based on the syntagmatic mechanism present in the linear unit. As will become apparent in Chapters 7 and 8, retrospective mechanisms play a much larger and, it could be argued, more important role in the IMDb corpus discourse than those related to prospection. In particular, discourse reflexivity features much more prominently in retrospective mechanisms than prospective mechanisms. It was therefore decided to limit this further sub-categorization to retrospective mechanisms only. As has already been seen in Chapter 2, Sinclair (1993/2004e) argues that the default retrospective linear mechanism present in written discourse structure is encapsulation, of which there are two types: deictic and logical encapsulation. He also proposes two 'exceptions' to this: verbal echo and overlay. An adaptation of these categories constitutes the categorization employed in this study. These are collectively termed *retrospective supplementary mechanisms*. The three categories of retrospective supplementary mechanisms are presented in Table 4.8 and will be discussed further below. [57]

Table 4.8: Categories of retrospective supplementary mechanism

Category	Description.
Encapsulation	The mechanism in the text whereby the linear unit which is acting as the text of the moment refers to the preceding linear unit by encapsulating it through labelling it and incorporating the label in

[57] In the IMDb corpus data, encapsulation can be found in both M and OI linear units. Overlay can only be found in M linear units and verbal echo only in OI linear units.

	the present linear unit. This labelling can range in terms of its explicitness.
Overlay	A point in the text where the linear unit which is acting as the text of the moment contains elements that act as a paraphrase or an antonym of the whole of the previous linear unit or of certain elements of the previous linear unit
Verbal echo	A point in the text where the linear unit which is acting as the text of the moment includes a *verbatim* repetition of the previous linear unit

These categories can be illustrated by referring back to the examples supplied in Steps 4 and 5 above. The examples in question will be repeated for ease of reference.

Following Moreno (2004) (see section 2.5.2) the term *encapsulation* in this study conflates Sinclair's categories of *logical* and *deictic encapsulation*. In the terms of this study, then, logical encapsulation is categorized as being a linear unit which contains an *implicit encapsulation* through an initial +ot– element, an ots element or an initial ⌊oi element. As we have seen in Chapter 2, Sinclair (1992) himself proposes a further type of implicit encapsulation in such acknowledgments as *yea* and *Mmm.* These will also be termed *implicit encapsulation* in this study contained within an OI linear unit.

Table 4.9: Levels of explicitness in encapsulation

Category	**Description**
Explicit encapsulation	Deictic acts, such as *these circumstances* or *these results* which explicitly indicate which part of the previous discourse is being encapsulated by a noun phrase with a determiner + referring noun
Fuzzy encapsulation	Occasions when the deictic element constitutes a pro-form, e.g.

	this, that, it, here etc.
Implicit encapsulation	Occasions when the reader is left unaided to work out which stretch of the previous text is being encapsulated, e.g. ot elements, descriptions of previous discourse acts in the text, referring nouns without a determiner, minimal acknowledgements (*Mmm*, *haha* etc.) and so on

An example of an explicit encapsulation can be seen in Example 4.28. The labelling noun *work* refers back to the whole of the Linear Unit in line 75 although the use of the definite article *the* in the phrase *the awesome work* makes the reference a little less obvious as to what is being referred to than the use of, say, the determiner *this* or *that.* Nonetheless, it is still considered an explicit encapsulation.

Example 4.28 (*Repeated*)

P	**L**	**Linear unit**	**Type**	**Retrospective supplementary mechanism**
N	74	Deniro can play a diverse range of roles and gives any movie an immediate sense of grandeur.. in much the same way as Al Pacino..	M	
	75	keep up the awesome work!!	MS	Encapsulation (implicit)

In Example 4.34, the encapsulation is achieved through the use of the pro-form *this* and is therefore termed a fuzzy encapsulation. As stated above, the use of a pro-form demonstrative for encapsulation is less explicit in its reference compared to a *that* + labelling noun. Here *this* would seem to refer to the thread in general, which at this point is only the linear unit in line 1.

Example 4.34 (*Repeated*)

P	L	Linear unit	Type	Retrospective supplementary mechanism
A	1	Petition for Dark Angel movie	M	
	2	**This isn't really a petition,**	**|M–**	**Encapsulation (fuzzy)**
	3	**just reply once** if you love "Dark Angel"	**+M**	

An example of implicit encapsulation occurring in an OI linear unit can be seen in the |OI linear unit in Example 4.45, where **Sigh*...* can be seen to be encapsulating and negatively evaluating the previous linear unit. Thus, the term *encapsulation* is distinguished in this study in terms of three degrees of explicitness (Moreno 2004) as seen in Table 4.9.

Example 4.45 (*Repeated*)

P	T	L	Linear unit	Type	Retrospective supplementary mechanism
A	1	1	Is it true he was bisexual?	M	
B	2	2	***Sigh*…**	**|OI**	**Encapsulation (implicit)**

In Example 4.35, the linear unit in line 2 contains the ot element *and* which acts as an implicit encapsulation. However, being a suspensive linear unit, *and* has two functions here. The first is what one would expect *and* to function as, i.e. to express the notion of 'in addition to the previously stated position (PSP)'. However, it also signals to the reader

that 'the upcoming linear unit will not be providing the contrast that you are expecting given the PSP'.

Example 4.35 (*Repeated*)

P	L	Linear unit	Type	Retrospective supplementary mechanism
A	1	It wasn't very pleasant	M–	
	2	and I'm glad there was liberal use of the black squares	M	Encapsulation (implicit)
	3	but it was hilarious.	+M	

As seen in section 2.5, Sinclair presents the 'exception' categories of *overlay* and *verbal echo* as distinct textual phenomena. The term *overlay*, according to Sinclair, refers to occasions where one linear unit can be interpreted as being a reformulation of the previous linear with a near synonymous meaning. *Verbal echo*, on the other hand, in Sinclair's terms, describes the repetition of part of the previous linear unit in a new context in order to provide a linguistic link to enable the writer to embark on a different topic. In the initial analysis for this study, this distinction proved to be extremely difficult to maintain. It was found that the majority of potential overlays in the IMDb corpus are in fact *selective* overlays (Sinclair 1993/2004e:95), either reformulating or taking *verbatim* only parts of the previous linear unit. However, these *selective* reformulations generally create a linear unit which has a connection with the previous statement but which cannot be termed synonymous. In Sinclair's terms, then, these would not be considered overlays as they do not mean the same. However, neither can they be considered to be a verbal

echo as it often is not a *verbatim* repetition. As noted in section 2.5, such difficulties with the delimitation of these types of linguistic mechanisms have been discussed particularly by Mauranen (2012). Due to such fuzziness in the delimitation of the two concepts, it was decided to resort to a clearer cut distinction between the two categories.

Overlay therefore is defined in this study as occasions when the linear unit which is acting as the text of the moment contains elements that act as a paraphrase or an approximate antonym of the whole or part of the previous linear unit. An important addition to the concept of overlay, therefore, is that an overlay may express something like the opposite meaning as well as the same meaning as the previous linear unit.

A clear example of this type of overlay can be seen in Example 4.38, where the phrase *their legacy* is repeated in line 26 from line 25 and the word *ridicule* acts as an antonym for the phrase is *not tarnished*. This overlay, then, provides a simple contradiction.

Example 4.38 (*Repeated*)

P	L	Linear unit	Type	Retrospective supplementary mechanism
I	25	"even great actors will work into their 80's and ridicule their legacy"	|OI	Verbal Echo
	26	**First off** their legacy is not tarnished	**MS**	**Overlay**

Example 4.39 illustrates another instance of an overlay in a suspensive linear unit where the overlay is being used in a reformulation which questions the right that the original poster has to pose the original question. The first two elements *Is It Time / to Retire* in

line 1 are reformulated in the first two elements of line 76 *what does deniro / have to retire*. However, this is hardly a neutral reformulation of the original question as the very rephrasing dismisses the validity of the original elicitation. Thus, the reformulation in the overlay is the prime indication that the linear unit in line 76 is suspensive.

Example 4.39 (*Repeated*)

P	**T**	**L**	**Linear unit**	**Type**	**Retrospective supplementary mechanism**
A	1	1	Is It Time to Retire Already?	M–	
O	25	76	what does deniro have to retire if he likes doing what he does:	|M–	Overlay
		77	acting	+M	

Another example of *overlay* can be seen in Example 4.37. In the suspensive linear unit in line 69, Poster M provides an antonym of the linear unit in line 68 by stating the opposite opinion. Therefore the opinion that De Niro should retire and is too old to perform certain roles is contrasted by the overlay *Deniro still the man!* It is true that such examples may be said to be an act of complex interpretation, as Sinclair admits, rather than a textually manifested phenomenon. However, it can be equally argued that there are certain linguistic signals of this contrast such as the verb *retire*, on the one hand, and the adverb *still* on the other, thus fulfilling the criterion stated above that it is only linguistically signalled links between linear units which would be categorized as being a qualified status M linear unit.

Example 4.37 (*Repeated*)

P	T	L	Linear unit	Type	Retrospective supplementary mechanism
A	22	68	he needs to / either / retire / or / need to / pick roles his age, / dammit.	+M	
M	23	69	Deniro still the man!	|M–	Overlay
		70	he can / pull out any roll / at his age	+M	

As indicated above, the notion of *verbal echo* in this study is distinct from that proposed by Sinclair. In this study, it is limited to examples where whole linear units are repeated *verbatim* such as in a citation using the [quote] [/quote] markup in initial positions in the turn. This facility has been described as conveying an 'illusion of adjacency' (Crystal 2001:141) in order to make the interaction more real. It is argued here that, rather than creating an illusion of adjacency, it allows the poster to drag a part of the previous discourse which has moved into the records of the autonomous plane back into the frame of focus so that the linear connection that the next linear unit makes with it is clear. As such, verbal echoes are not considered to be message-oriented as they are simply repeating part of the previous discourse in order to clarify to which part the subsequent part of the turn is related. They are therefore categorized as OI linear units. There are several examples of this in the *De Niro* thread, seen particularly when the poster uses the [quote] [/quote] markup facility. Typically, this manifests itself in an apparent internal dialogue within the turn in which the poster responds to the quotation. However, in the linear analysis in this study, the verbal echo will be considered to be a suspensive |OI linear unit and the subsequent linear unit will be seen to be engaging with the original linear unit in the previous turn. For instance, in Example 4.38, the linear unit in line 25 is

a quotation from line 9. This is regarded as suspensive |OI linear unit as the simple repetition of a linear unit by another discourse participant is not expected as normal practice in discourse exchanges. The linear unit in line 26 is therefore seen as a retro inter-turn |M linear unit as it contradicts the opinion provided in line 9.[58]

Example 4.38 (Repeated)

P	T	L	Linear unit	Type	Retrospective supplementary mechanism
E	5	9	like a boxer even great actors will work into their 80's and ridicule their legacy,	+M	
I	9	25	"even great actors will work into their 80's and ridicule their legacy"	|OI	Verbal echo
		26	**First off** their legacy is not tarnished	**|M–**	

4.3.3.8 Points of connection in linear units with retro inter-turn orientation

In order to use the system of analysis successfully in turn-initial linear units it is important to establish what will be termed in this study, *the points of connection*, i.e. the exact linear unit in the previous turn from which the second linear unit follows. The coherence of online message board discourse differs significantly from the monologic written discourse on which the analytic system for written discourse (Sinclair 1993/2004e) is based. The fact that online message board discourse is dialogic, or, as is the case in the *De Niro* thread, polylogic, means that any inter-turn sequence is almost always the

[58] As such, |OI linear units containing a verbal echo are considered to be one referential action. Therefore, the whole of the text within the [quote] [/quote] markup is classified as one linear unit however many linear units they constituted in the original post.

product of two or more contributors just as it is in spoken dialogue. However, unlike face-to face spoken interaction, the fact that online message board discourse is also asynchronous means that, in responding to a previous turn, the poster is able to choose which linear unit in the previous turn in the nest structure to follow on from (see section 4.2.1). The point of connection may be the first linear unit of the turn and the final linear unit of the turn immediately previous in the nest structure of the thread. However, it may also be a linear unit from earlier in the turn. This means that strict linearity need not be followed in the first linear unit of a turn if the point of connection is clearly previous to the last linear unit. This has implications for the categorization of the first linear unit of a turn as will be illustrated by looking at the first ten turns of the *De Niro* thread in this respect. These are reproduced in Example 4.46. The nest structure of these ten turns can be seen in Figure 4.3.

Example 4.46: The first ten turns of the *De Niro* thread

P	T	L	Linear Unit	Type
A	1	1.	Is It Time / To Retire / Already?	M–
		2.	C'mon / man, if he keeps on making / mobster movies / and / bubble gum movies / like / Meet The Parents / or / Meet The Fockers,/ then / that's just whack.	|M
		3.	**Just tell me, / people,** is it Time for him / to Retire?	M–
		4.	I wanna know.	|OI
B	2	5.	NO!	+M
C	3	6.	He has done / some pretty crap films / in the last 5 years.	+M
D	4	7.	De Niro will revive his career, / and / amke you guys / **eat those words.**	**|M**
		8.	**To answer your question /** NO	MS
E	5	9.	like a boxer / even great actors / will work into their 80's / and / ridicule their legacy,	+M
		10.	brando-olivier / they have all done it	MS
F	6	11.	yeah i couldnt agree with you more / buddy.	MS–

		12.	this guy / hasn't made a good film / in along time.	+M
		13.	**i mean,/** he made those two films	MS–
		14.	godsend / and / hide and seek.	+M–
		15.	they were both / the same goddamn movie.	+M
		16.	i think / he should / follow the example of / good actors / like / hugh grant / and / jude law/ and / choose his roles / a bit better	M
G	7	17.	[quote] i think he should follow the example of good actors like hugh grant and jude law and choose his roles a bit better [/quote]	|OI
		18.	**ROTFLMAO!!!**	**|OI**
H	8	19.	hey, /hugh grant and jude law / are great actors!	|M–
		20.	they made some fantastic movies / like	+M–
		21.	euh…/ hmm… / ok / **wait…**	**|OI**
		22.	hmm... / **no wait** / i got one...	**|OI**
		23.	hmm / no....	|OI
		24.	ROFL	OIS
I	9	25.	"even great actors will work into their 80's and ridicule their legacy"	|OI
		26.	**First off** / their legacy is not tarnished	**|M–**
		27.	because /once they die / the bad films / will be forgotten / anyway.	+M
		28.	**Second**, / we need someone / to play the parts / for older people...	**MS–**
		29.	or /do they not even exist / to you?	+M–
E	10	30.	of course / they ruin their legacy,	|M–
		31.	when they ham up / in rubbish films / they make themselves / look like fools,	+M
		32.	when Olivier did Inchon-Jazz Singer / etc / he was absolutely / making a fool of himself / and / looked ridiculous.	MS
		33.	so / **to say** / you dont ruin your legacy / is rubbish	**MS**
		34.	because / you do.	MS
		35.	Pacino selects / what he does / v carefully	M

From Example 4.46, we can see that Turn 1 ends with a pro inter-turn M– linear unit (if we ignore the final |OI linear unit) in which Poster A asks the question s/he wants others to respond to. Three Posters: Posters B, C and E respond directly to this elicitation in

Turns 2, 3 and 5 respectively. These three responses are all compliant as can be seen by the fact that they begin with a +M linear unit thus forming an inter-turn M– /+M sequence. However, strictly speaking they have not responded to Turn 1. Rather, they have responded to the linear unit in line 3.

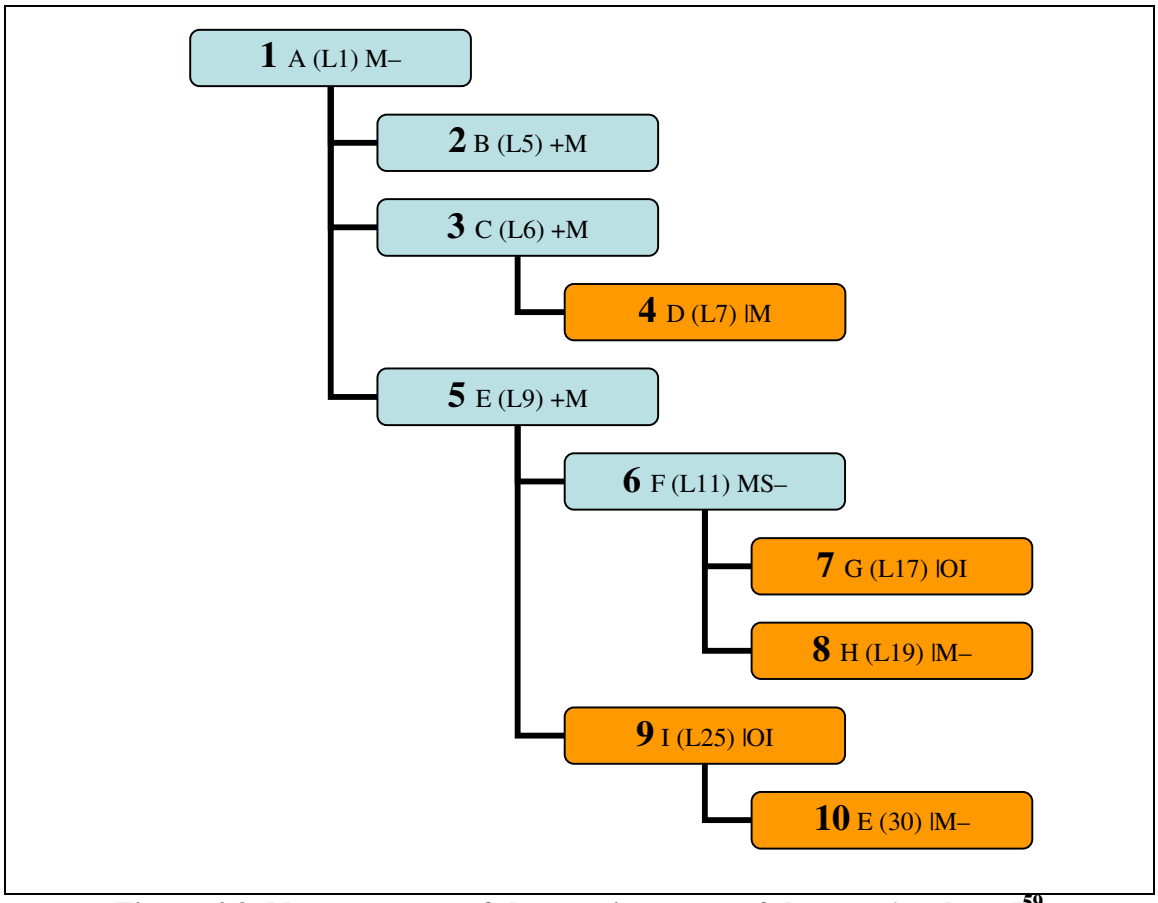


Figure 4.3: Nest structure of the opening turns of the *De Niro* thread[59]

In Turn 6, it is unclear what exactly Poster F agrees with in Turn 5. It would seem probable that s/he does not simply agree with the examples provided in line 10 in the linear unit *brando-olivier they have all done it*. However, it remains unclear. Ultimately

[59] In this diagram each lozenge gives the following information: Turn number; Poster; Line number; the first linear unit of the turn (except the first turn in which the salient linear unit of the point of connection is given). A lighter shaded lozenge signifies compliant first linear unit; a darker shaded lozenge suspensive first linear unit.

this vagueness does not affect the analysis as *you* in line 11 is considered to be an example of *fuzzy encapsulation*.

It should be noted that after the establishment of the point of connection between turns, the rest of the present turn is treated strictly based on linear sequentiality. For example, in Turn 4 in line 8, the poster signals that s/he is about to respond to the elicitation in line 3. However, in a linear analysis, this linear unit is considered to be primarily signalling that there is some sort of linear relation, not between itself and line 3 but simply with the immediately previous linear unit in line 7. In linear terms, then, it signals that an overlay, i.e. stating the previous linear unit in the terms of the original elicitation, is coming up.

Thus, it is only the first linear unit of each turn which may not follow a strict linear sequence. These are signalled by using the [quote] [/quote] markup facility, as we see in Turn 7 in line 17 or by the use of quotation marks as we see in line 25. It should also be noted that establishing a point of connection can sometimes be a matter of interpretation as many of the points of connection are made without the use of the [quote] [/quote] markup facility. For instance, the first linear unit in Turn 10 in line 30 may be interpreted to be establishing a point of connection with the linear unit in either line 26 or 27. In the end it was felt line 26 was most likely due to the strong presence of an overlay through antonyms.

Having provided an overview of both the systems of analysis of element and linear unit relations, in the next section issues related to the classification of elements by discourse reflexivity will be discussed.

4.4 Discourse reflexivity

As stated in Chapter 1, discourse reflexivity, as employed in this study, is defined as:

> A property evident in certain elements whereby there exists an explicit reference to the present discourse through a reference to one of the following: the text as text; a discourse event; a discourse act; the writer as text-constructor; the reader as text-decoder; a place in the text or the time or manner of the discourse act.

As such the concept of discourse reflexivity broadly aligns itself with the 'reflexive model' of metadiscourse as described in Chapter 3 and seen in Mauranen (1993a, 2007, 2010, 2012) and Ädel (2006, 2008, 2010), although the term *metadiscourse* is for the most part not employed in this study. As will be seen, it is, however, a stricter interpretation of explicitness than seen in previous studies with instances of low explicitness (Mauranen 1993a) excluded as are instances of *writer-reader* or *audience interaction* (Ädel 2006, 2010) that do not clearly contain an explicit reference to the discourse as discourse.

A decision was also required as to which of the two basic analytical units – the linear unit or the element – was to be categorized as being DR or non-DR in this study. According to

Ädel (2006), there are strong arguments for the use of small reproducible units in the description of discourse reflexivity in order to provide the necessary precision for a fine-grained study. Combined with this is evidence that discourse reflexivity has a tendency to occur in short fixed and semi-fixed expressions (Mauranen 2010), which would very often coincide with an element rather than a linear unit. It was therefore decided that the element would be used as the unit for discourse reflexivity in this study.

Discourse reflexivity, therefore, is envisaged as being a quality that certain elements possess. The quality that these elements have in common, as stated above, is that they refer explicitly to the present discourse. This may include elements from any of the three main element categories: m, oi or ot elements. These will be illustrated below.

DR m elements are therefore conceived as being message-oriented elements which lead to the incrementation of shared experience and which contain an explicit reference to at least one of the six categories listed in the definition above (see Table 4.10).

Table 4.10: Categories of discourse reflexive elements

	m elements	**oi elements**	**ot elements**
Text as text	*that's HER words; this thread*	*To answer your question*	*Just on a little side note*
Discourse event	*your statement; an endless off-topic discussion*	*To answer your question; a question to you*	*in conclusion; for example; for instance*
Discourse act	*let's take a look at; you said; to say*	*I mean; wait; tell me*	*by contrast; to repeat;, to conclude*
Discourse	*you said; I'd like to*	*I mean; as I say; tell*	~

participant	*mention; the OP needs glasses*[60]	*me; OP*	
Place in the text	*Nothing special here*	*as mentioned above*	~
Time of the discourse act	*already*	*again*	*first off; second; finally; lastly*
Manner of the discourse act	*I'll tell you plainly, just mindlessly respond.*	*frankly speaking, seriously*	*in short*

The elements containing a reference to the text as text include what Francis (1994) terms text nouns (*phrase*, *section*, *paragraph* etc.), language activity nouns (*contrast*, *example*, *debate* etc.) and illocutionary nouns (*accusation*, *comment*, *suggestion* etc.), as seen in Chapter 2. The 'present discourse', referred to in the above definition, includes other posters' turns within the same thread, thus following Mauranen (2012), and other threads within the same message boards just as Ädel (2010:76) includes different lectures from the same series in her model of metadiscourse.

As regards the place in the text, neutral deictic pro-forms (Ribera 2007), i.e. *that*, *this*, *these*, *those* are not considered to be explicit enough to be considered in this study to be DR (unlike Schiffrin 1980) whereas the locative deictic pro-form *here* is. This judgment was made based on the fact that *here* would seem to carry more specific reference related to its surrounding, whether it be in relation to the whole thread as in Example 4.47 or to a specifically prospected linear unit in an act of advance labelling (Tadros 1994) as in Example 4.48, compared to the vague reference seen in *this* and *that*.

[60] OP is an acronym which stands for *Original Poster* (*www.urbandictionary.com*)

Example 4.47[61]

P	L	Element	Element type
A	1	**ive not once**	**m–**
	2	**mentioned about**	**+m–**
	3	deniro	+m
	4	**in post**	**ms**
	5	**here**	**ms**
	6	**in this thread**	**ms**

Example 4.48[62]

P	L	Element	Element type
A	1	**Here's my Christmas list:**	**m**
	2	**1:**	**+ot–**
	3	Direct me	m–
	4	a couple of Westerns	+m
	5	**2:**	**+ot–**
	6	Make a final Dirty Harry	m

DR oi elements, on the other hand, are elements whose function is the management of the interaction and which contain an explicit reference to one of the six categories (see Table 4.10). DR ot elements are defined as elements whose function it is to manage the internal coherence of the text, i.e. those which display internal relations (see section 3.3.3).

As can be seen in Table 4.10, there was no evidence from the IMDb corpus data that ot elements make explicit reference to the discourse participants. Instead, they remain

[61] This example is from Thread 001, in which the relative merits of American actors Al Pacino and Robert de Niro are discussed.
[62] The example is from Thread 006 about the American actor and director Clint Eastwood in which the poster rather whimsically writes a Christmas wish list as if to Eastwood.

impersonal in their linking of parts of the text. There is similarly no evidence of ot elements including an explicit reference to a place in the text.

Having presented the system of analysis for element relations and linear unit relations as well as the definition of discourse reflexivity, the next section will describe the analytical procedures carried out to answer the research questions of this study.

4.5 Analytical procedures

4.5.1 Features of the *UAM CorpusTool*

The tool for analysis which was employed in this study was mainly the *UAM CorpusTool* (version 2.8) (see http://www.wagsoft.com/CorpusTool) (O'Donnell 2009). The *UAM CorpusTool* is principally a semi-automated tool designed for the annotation of corpus data. It has been employed in a variety of research areas, mainly in the areas of learner errors (e.g. Lozano 2009) and Systemic Functional Grammar (SFG) (e.g. Moore 2011). As Sinclair and Mauranen (2006:156) comment, LUG can be laid out in terms of a series of binary choices as if it is a paradigmatic system in the style of SFG, though the choices (e.g. between m elements and o elements) cannot be considered to be genuine paradigmatic choices made by the speaker/writer as conceived in SFG. However, Sinclair and Mauranen (2006:8) also comment that even though LUG is syntagmatically oriented, the sub-categories of elements (m–, +m -, ms etc.) are essentially systematic in that 'they are small sets of mutually exclusive choices'. The fact that the categorization of LUG is often binary or at most a choice of a limited number of categories means that even though the *UAM CorpusTool* is designed for use in SFG-style paradigmatic analysis, it can be

employed for the purposes of this study, albeit by ignoring some of the terminology employed by the tool, as will be seen below.

For the purposes of this study, the *UAM CorpusTool* has three features which made it particularly useful. Firstly, it is possible to design a particular annotation scheme or schemes for the study. This was particularly important given the innovative nature of the models being proposed. Secondly, the *UAM CorpusTool* allows the researcher to individually decide on the length of each segment. This is particularly useful for the segmentalization of non-standard or varied segments as is the case with the element and the linear unit in this study. Thirdly, the *UAM CorpusTool* allows for the annotation of the same text using different annotation schemes, or *layers* to use the term employed by the *UAM CorpusTool*. In this study therefore the *UAM CorpusTool* permitted the search for instances of a certain element type in all of the corpus or in a certain type of linear unit. For instance, an *in* search may have been a search for *oi elements* ***in*** *+M linear units with inter-turn orientation.* A *containing* search could be a search for *|M linear units* ***containing*** *discourse reflexive elements*.

As will be seen in Chapters 5 to 7, such searches of the *UAM CorpusTool* within one annotation scheme or across both annotation schemes form the core of the results produced for this study.

4.5.2 The two annotation schemes

In this study, two annotation schemes (or *layers*) were developed: one for the analysis of element relations (Figure 4.4) and the second for the analysis of linear unit relations (Figure 4.7). For a step-by-step set of instructions as to how to create the annotation scheme for element relations and linear unit relations see Appendix 4.5.

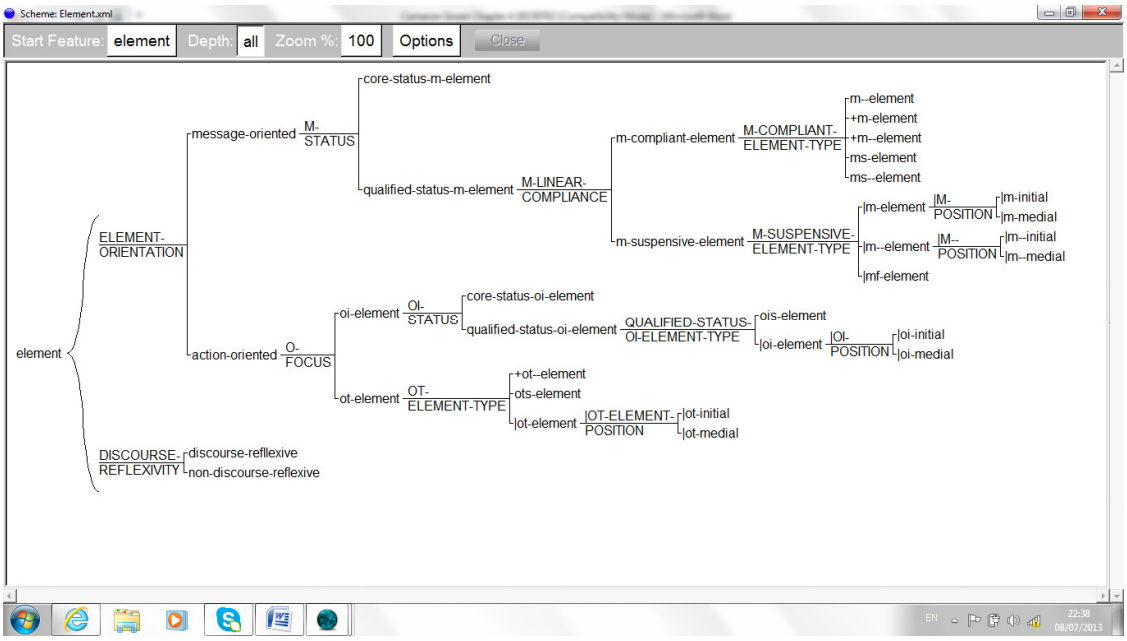


Figure 4.4: *UAM CorpusTool* annotation scheme: element relations

The first stage of the annotation is to decide the extension of the element. Taking the *De Niro* thread as an example, the first sentence is the subject line *Is it Time to Retire Already*. Figure 4.5 shows the designation of the first element *Is it time* underlined and shaded in the top left-hand corner of the screen. The researcher is then asked to make a decision from a series of choices based on the annotation scheme. The first decision, as can be seen in Figure 4.5 in this case, therefore, is whether the element is message-oriented or action-oriented.

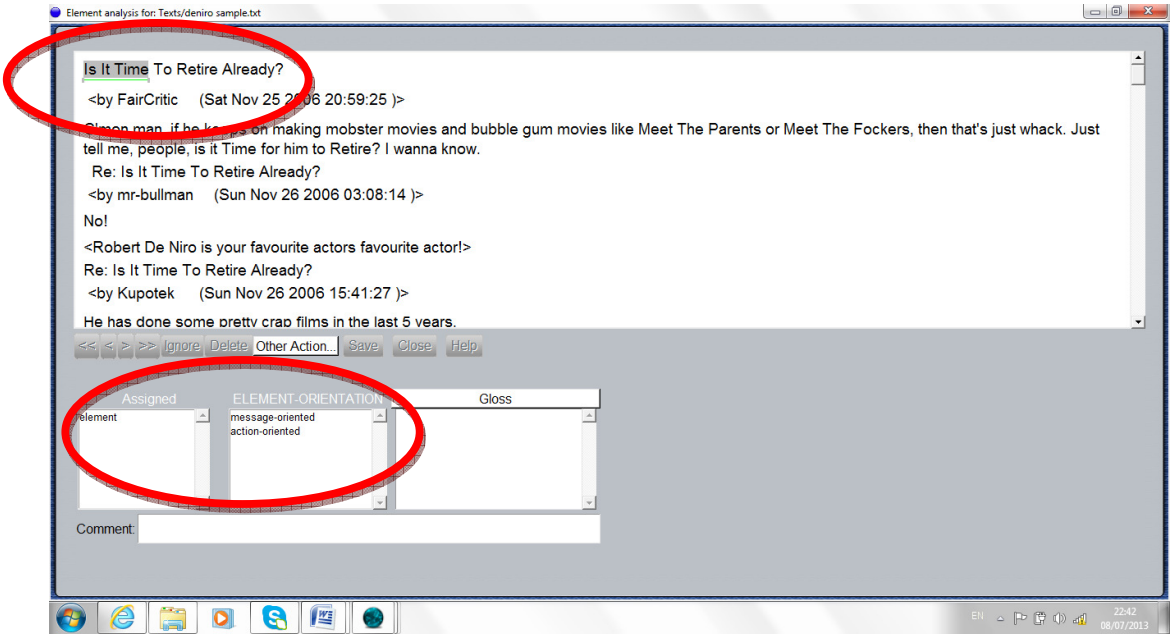


Figure 4.5: Assignation of first element of thread

The annotation scheme, then, reflects the seven steps described in section 4.3.1 as regards element relations and discourse reflexivity. It can be summarized as having two fundamental separate binary decisions, seen within the curly brackets on the left-hand side in Figure 4.4. These provide two branches of decisions to be made in order to complete the annotation.

The uppermost branch (as seen in Figure 4.4) relates to element orientation, i.e., whether the element is message or action-oriented. Once this decision has been made, a series of further decisions are asked to provide a finer categorization of the element. These will be outlined below. The second fundamental decision to be made is whether the element is DR or non-DR. No further decision is made in this branch.

In element orientation, if message-oriented elements have been selected then the status of the m element is next decided upon, i.e. whether the m element is core or qualified status. If it is core, then, no further decisions are needed. If it is qualified, then, the next decision to be made is whether the qualified m element is compliant or suspensive. Each of these then lead to further decisions to be made until the right-hand extreme of each branch is reached. If the decision is that the element is an action-oriented element as opposed to message-oriented, then, a decision regarding its focus has to be made, i.e. whether it is an oi or ot element. Again, at each juncture a selection is made from a choice of two or three options until the end of the branch on the right-hand side is reached. At the end of this decision making process the element is annotated as seen in Figure 4.6. In this case the element is annotated as being: message-oriented; qualified status m element; m suspensive element; |m– element; |m– initial; and non-discourse reflexive.

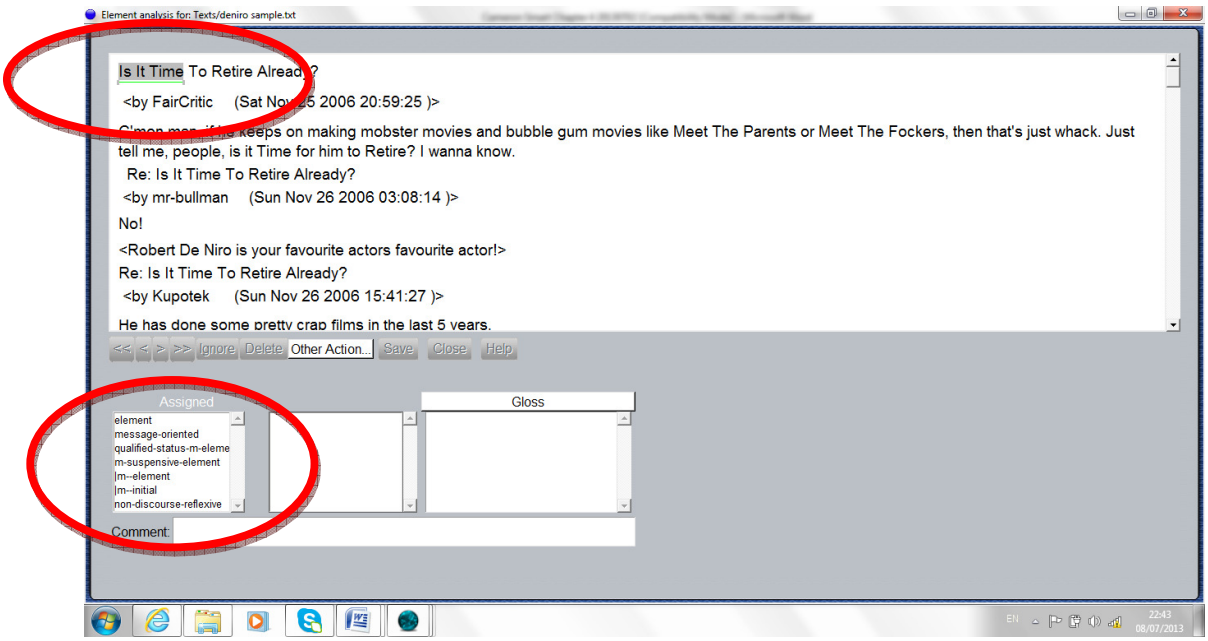


Figure 4.6: Assignation of first linear unit of thread

Once a thread has been completed in this way, the same thread is annotated employing the other annotation scheme, i.e. that of linear unit relations. For this layer the annotation scheme seen in Figure 4.7 is followed.

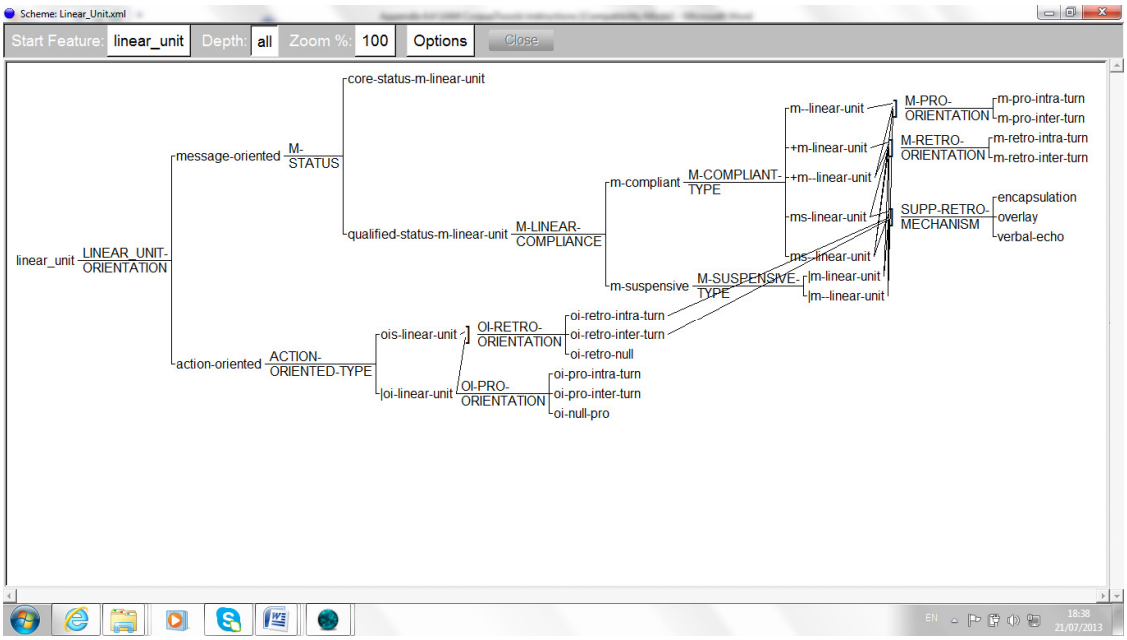


Figure 4.7: *UAM CorpusTool* annotation scheme: linear unit relations

For linear unit relations, the first decision made is the extension of the linear unit. This decision is entirely based on Step 1 as seen in section 4.3.3.2. After deciding on the length of the linear unit a series of choices are made similar to those described above for element relations analysis. As can been seen in Figure 4.7, there are similarly two branches in the linear unit relations annotation scheme. Hence the first decision to be made relates to whether the linear unit is message or action-oriented (see Figure 4.8). A series of choices are then presented as the annotation proceeds along the chosen branch in a similar fashion to the element relations annotation scheme above.

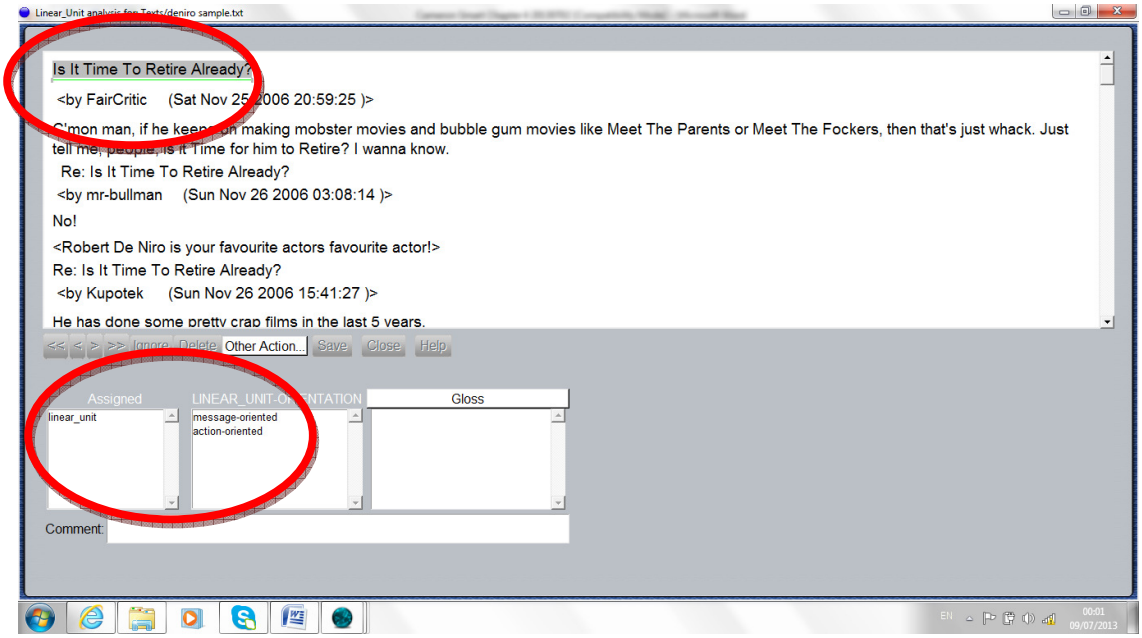


Figure 4.8: Assignation of first linear unit of thread

However, it can be seen in Figure 4.7 that the annotation scheme for linear unit relations is slightly more complex for linear unit relations than that for element relations approaching the end of the branches on the right-hand side of the screen. This is due to the fact that certain sub-categorizations do not apply only to one particular type of linear unit. For instance, M pro orientation, i.e. whether incomplete M linear units prospect completion within the turn (pro intra-turn) or prospect completion from another turn (pro inter-turn), not only applies to M– linear units but also to +M– linear units, MS– linear units; and |M– linear units. In order to convey this choice the *Change Entry Condition Function* can be used. In this function, the number of terms related to the entry condition can be selected, thus specifying to which categories the sub-categorization applies. In this case, there are four *terms*: M– linear units, +M– linear units, MS– linear units and |M– linear units as can be seen in Figure 4.9.

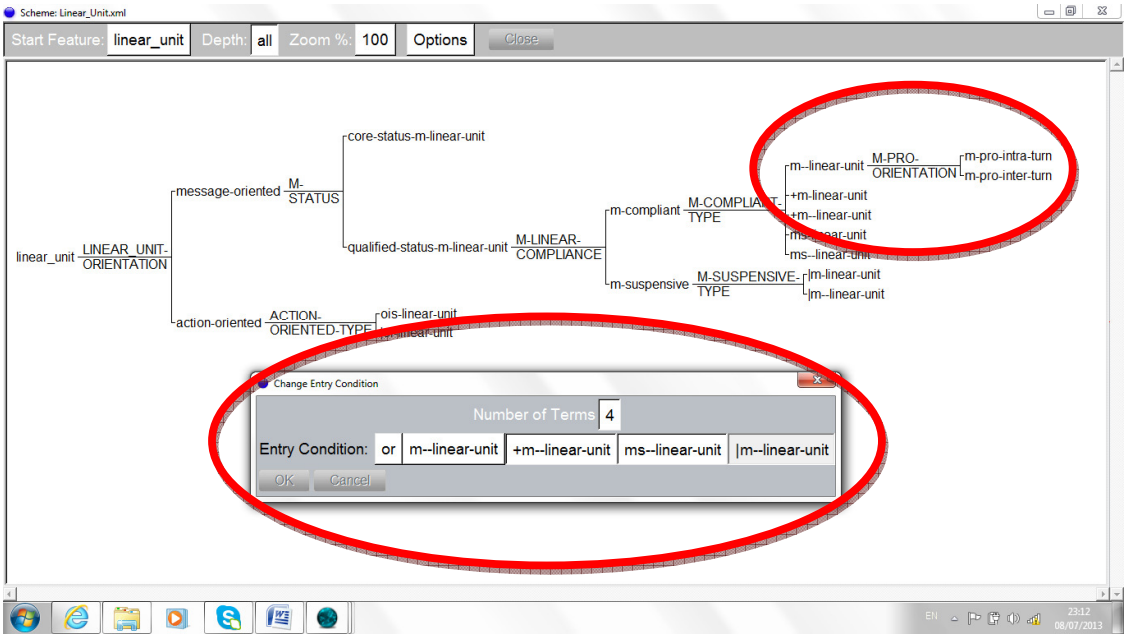


Figure 4.9: Entry condition for M pro orientation system

This produces the annotation scheme seen in Figure 4.10. Hence, only when one of these four linear unit types are reached does a decision have to be made regarding the M pro orientation of the linear unit. The occasions where the change entry condition function is used are detailed in Appendix 4.5.

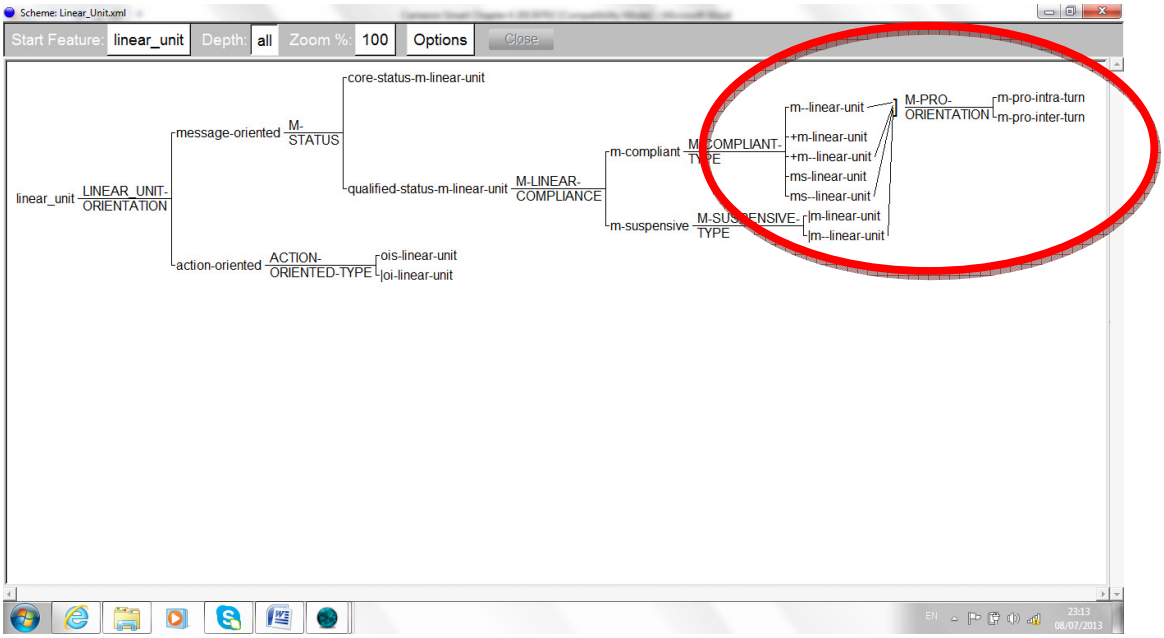


Figure 4.10: Annotation scheme of pro orientation with four-term entry condition

Thus, following the annotation scheme, the first linear unit of the *De Niro* thread is categorized as seen in Figure 4.11. The first linear unit is therefore: message oriented; qualified status; compliant; M– linear unit; with pro inter-turn orientation.

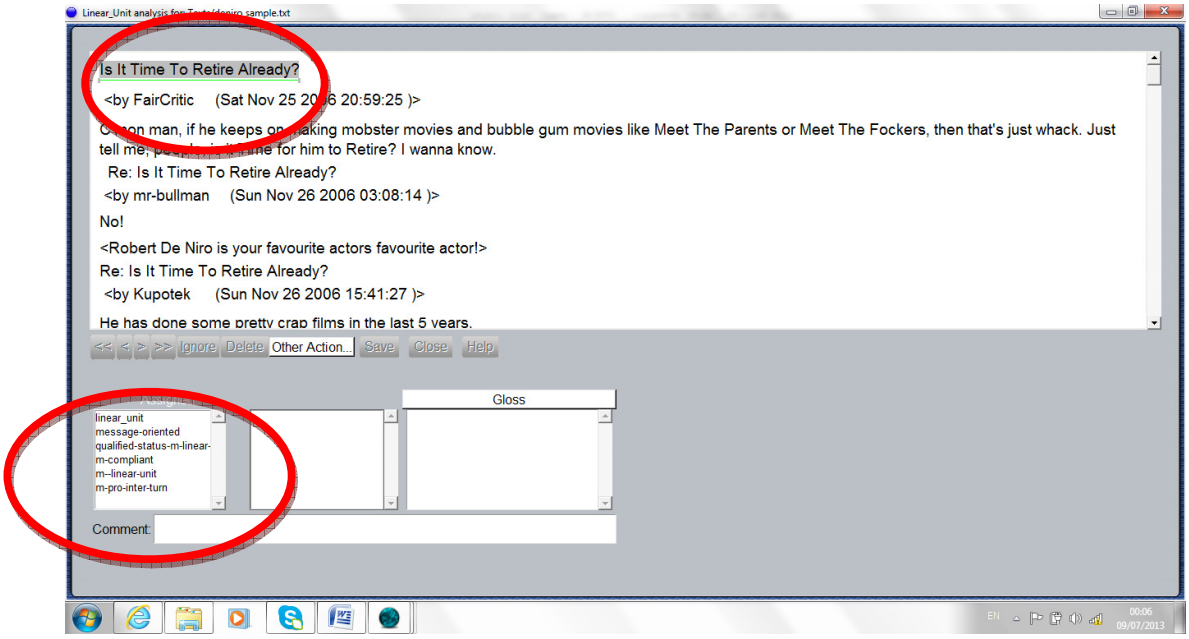


Figure 4.11: Assignation of first linear unit of thread

After the completion of the annotation of the linear unit, the same procedure is carried out for the next linear unit and so on. The same procedure was then followed for all other 70 threads in the IMDb corpus.

4.5.3 Analysis of Data

4.5.3.1 Frequencies and relative distribution

Three main statistical procedures were employed in answering the research questions of this study from the annotated data gleaned from the procedures described above. These will be described below.

The crudest statistical information used in this study is that of frequency lists. These come in two forms: firstly, the frequency and relative distribution of categories and subcategories of the annotation schemes as described above. This can be seen in use throughout Chapters 5, 6 and 7. Secondly, Frequency WordLists from *WordSmith Tools* (Scott 1996) were also employed. In order to do this, a search for a particular category or subcategory was performed using *UAM CorpusTool*. This produced a list of all examples of the category in question, which could then be extrapolated and saved as a text-only file in order to carry out a Frequency WordLists using *WordSmith Tools.* This can particularly be seen in use in Chapter 5. As this is not a study in which two corpora were to be compared with each other, frequency lists were generally used in conjunction with one or both of the other statistical operations described below.

4.5.3.2 Split-corpus calculation

The central question to be answered in this study is whether DR elements are distributed across the categories and subcategories of the two systems of analysis in the same way as non-DR elements are or whether the number of DR elements in a particular category or subcategory is higher or lower than would be expected if they were distributed in the same way as non-DR elements.

If we simply examine the frequency totals of DR elements and non-DR elements in a particular subcategory, (e.g. 13 DR oi elements and 97 non-DR oi elements), there is no way of ascertaining whether the number of DR elements is higher or lower or the same within this category as what would be expected if they were distributed in the same way as non-DR elements. In order to calculate this, a calculation of statistical significance is required. This was done by firstly dividing the corpus, in effect, into two subcorpora: a DR element corpus and a non-DR element corpus. Such a split-corpus approach has previously been established in several studies (Mahlberg and O'Donnell 2008; O'Donnell et al 2012). The procedures employed here is an extension of such methodology.

One of the most frequently used calculations of statistical significance is the Pearson's chi-squared test (χ^2) (Brown 1988; McEnery and Wilson 2001). However, it has been found that Pearson's chi-squared test may be unreliable when dealing with very low frequencies of occurrence and over-estimates significance in corpora of very different sizes (Dunning 1993). Both of these were the case in this study since the IMDb corpus is relatively small and the total number of DR elements was obviously going to be much lower than the number of non-DR elements. For that reason the Log Likelihood Ratio (G^2)

(Dunning, 1993; Rayson and Garside 2000; Leech et al 2001) was employed in the split-corpus calculations in order to calculate whether the difference between the actual observed frequency of DR elements in a particular category is statistically significant when compared to what the expected frequency would be if they were distributed in the same way as non-DR elements in the corpus.

If no significant difference was found between the observed and expected frequencies then it was concluded that DR elements were simply acting like their non-DR counterparts. If a significant difference was found, it was judged that the DR elements are over-represented or under-represented in that particular category. It should be noted that by over and under-representation we are not attaching any evaluation to the frequency. It is, rather, simply a statistical expression of the presence of significant difference, in the same way as the terms *overuse* and *underuse* has been used other corpus-based studies (e.g. Altenberg and Tapper 1998).

The p value in such calculations was set at $p < 0.001$ for an over/under-representation of high significance meaning a critical value = 10.83 (indicated in the tables in this study by **). Additionally, the p value was set at $p < 0.01$ for an over/under-representation of low significance meaning a critical value = 6.63 (indicated in the tables in this study by *) (Rayson et al 2004). However, according to Rayson et al (2004) with a frequency of less than five, the p-value should be set at $p < 0.01$ and the critical value should be set at 15.13. Therefore, in this study, for cases of frequency less than five, a critical value was set at 15.13 and even with this higher critical value it was still only classified as an over/under-representation of low significance.

4.5.3.3 KeyWord analysis

KeyWord analysis was employed (Scott 1997, 2001) using *Wordsmith Tools* by using a similar split-corpus approach to the above in order to ascertain which words in each category were more salient. In this approach, two WordLists are produced: the first the category under investigation, the second a WordList of the opposite of the feature under investigation. For example, if DR m elements were under investigation, then, two WordLists, one of DR m elements themselves and one of non-DR m elements would be drawn up. These WordLists were then used to produce a KeyWord list in order to find the most salient words in the first subcorpus. This, then, provided the basis for which elements to investigate. This methodology, again, is similar to the methodology used by O'Donnell et al (2012) to investigate what they call 'intra-textual keyness'. The significance of the 'keyness' of the word in question is calculated employing log likelihood calculation as part of *Wordsmith Tools* (Scott 2001).

There are two *p* values used in the KeyWord searches in Chapters 5 and 6. In the Keyword searches in Tables 5.6, 5.17 and 6.13 the *Wordsmith Tools* default $p < 0.000001$ is used (1E-6 in scientific notation). In the KeyWord searches in Tables 5.23, 5.29, 6.16 and 6.20, the critical *p* value was lowered to $p < 0.0001$ (or 1E-4) in order to illustrate more instances as the frequency of individual words is low. Any conclusions drawn from these searches are consequently more tentative.

4.6 Research questions

As stated in Chapter 1, the overarching research question of this study is as follows:

> *Do DR elements behave in the same or in different ways in the linear structure of IMDb corpus discourse when compared to their non-DR counterparts?*

This question will firstly be divided into two broad scopes of analysis: element relations analysis and linear unit relations analysis. As also stated in Chapter 1, this produces two overarching questions:

1. *Do **DR elements** behave in the same or in different ways when compared to their non-DR counterparts in terms of the relations between elements in IMDb corpus discourse?*
2. *Do **linear units which contain DR elements** behave in the same or in different ways when compared to linear units which do not contain DR elements in terms of the relations between linear units in IMDb corpus discourse?*

The categorization of elements and linear units in the present system of analysis as detailed in this chapter is based on two factors. The first of these is what Sinclair and Mauranen call the orientatiation of the element or the linear unit, i.e. if it is message or organizationally-oriented. The second factor is the status of the element or linear unit in the sequential syntagmatic relations of the discourse, i.e. how each element or linear unit is related to the previous and upcoming one. Thus, following the second of these factors,

the relations between elements in a linear description can be quantified by measuring the relative frequencies of the main element categories and subcategories. Similarly, the relations between linear units can be quantified through measuring the frequency of the different linear unit categories and subcategories.

The two overarching questions above can therefore be expressed in terms of the following seven research questions, each of which contains further research sub-questions, which will be detailed in the relevant sections of the Chapters 5–7. The first four of these questions deal with element relations: firstly, in general; and then in terms of the three element categories (m, oi and ot elements). The fifth, sixth and seventh questions deal with linear unit relations: firstly, in general; and then in terms of the two main linear unit categories (M and OI linear units):

Research Question 1: *How are elements in general - and DR elements in particular - distributed in the IMDb corpus discourse?*

Research Question 2: *How are m elements in general - and DR m elements in particular - distributed in the IMDb corpus discourse?*

Research Question 3: *How are oi elements in general - and DR oi elements in particular - distributed in the IMDb corpus discourse?*

Research Question 4: *How are ot elements in general - and DR ot elements in particular - distributed in the IMDb corpus discourse?*

Research Question 5: *How are linear units in general - and linear units that contain DR elements in particular - distributed in the IMDb corpus discourse?*

Research Question 6: *How are M linear units in general - and M linear units that contain DR elements in particular - distributed in the IMDb corpus discourse?*

Research Question 7: *How are OI linear units in general - and OI linear units that contain a DR element in particular - distributed in the IMDb corpus discourse?*

The findings relating to the first of these questions will be presented in Chapter 5. The aim of Chapter 5 is therefore to establish the characteristics of DR elements in the three main categories of elements: m, oi and ot elements. Research questions 2–4 will be answered in Chapter 6. Chapter 6 will therefore describe the characteristics of DR elements in the subcategories of elements as well as their role within the linear unit and between linear units. Research questions 5–7 will be answered in Chapter 7 through a detailed presentation of the findings concerning linear unit relations.

4.7 Chapter summary

This chapter may be considered to be the centerpiece of this study in that both the systems of analysis for element and linear unit relations as well as a definition of discourse reflexivity have been presented. In the system of analysis for element relations it has been shown that a system based on LUG was employed here. However two important embellishments of the original model were also presented. Firstly o elements are sub-categorized in this model. Secondly, the concept of suspension has been introduced to the model for all three main types of elements. As will be seen in Chapters

5 and 6, this inclusion provides one of the central findings of this study as regards the distribution of DR elements.

It has also been shown that this study proposes to unify the system of analysis of linear unit relations with that of element relations by employing an adaptation of the same codification as LUG used for elements. Thus, prospection, completion and encapsulation are expressed in terms of M and OI linear units and their subcategories. As well as this, the key concept of suspension is also incorporated into the model of analysis for linear unit relations. Again, this will prove to be crucial in terms of the findings as regards the distribution of discourse reflexive elements as will be seen in Chapters 7 and 8. Finally, a clearly delimited definition of discourse reflexivity has also been presented.

Also described in this chapter was how the systems of analysis were converted into workable annotation schemes for the *UAM CorpusTool*. Finally, the statistical measures employed in the results section were also introduced with the Log Likelihood Ratio and KeyWord search being the most important of these.

In the next chapter the results for the system of analysis of element relations will be presented.

CHAPTER 5

DISCOURSE REFLEXVITY IN ELEMENTS IN THE IMDB CORPUS DISCOURSE

5.1 Overview

This chapter will be dedicated to answering the first research question as stated in section 4.6:

Research Question 1: *How are elements in general – and discourse reflexive (DR) elements in particular – distributed in the IMDb corpus discourse?*

In order to answer this question, the frequency of the three main categories of elements: m, oi and ot elements, will be presented in section 5.2.1. Having established the relative frequency of each category of elements in general, each of these will then be re-examined in section 5.2.2, where the frequency of DR and non-DR elements will be compared in order to find the points in the IMDb corpus discourse where significant differences in the behaviour of DR elements occur compared to their non-DR counterparts. The lexical characteristics of each of the three main categories will be presented in section 5.3. In the case of o elements, their non-structural discourse functions will also be presented. Finally, in section 5.4, the lexical characteristics of DR elements within each of the three main categories will be presented as well as the non-structural discourse functions of DR o elements.

Linear analysis will be employed in this chapter as a means of establishing the boundaries of each element and in the categorization of the element in terms of the three main categories. However, in order to describe the lexical characteristics of the DR elements within each category dictionary definitions of words and phrases within the element will be drawn upon. In the case of o elements their non-structural discourse function will also be presented using an adaptation of Hyland's (2005) categorization of metadiscourse. This will then provide the basis for the further exploration of DR elements in the subcategories and their role within the linear analysis of this study in Chapter 6.

As described in Chapter 4, the research in this study is based on the annotation of the IMDb corpus using the *UAM CorpusTool*. The corpus comprises a total at 15,102 elements, which are made up of 41,195 words. This means that the average element length is 2.73 words.

5.2 Three main element categories

5.2.1 Relative distribution of m, oi and ot elements

Research Sub-question 1a*: What are the relative frequencies of the three main categories of elements: m, oi and ot elements?*

Based on the figures in Table 5.1, in the most general of terms, the IMDb corpus discourse can be characterized as being message-oriented, with over 76% of all elements categorized as being m elements. These message elements are supported by a large

number of ot elements, at 15.93% of all elements. This figure is almost twice the number of oi elements, which represent 7.58% of all elements.

Table 5.1: Elements in the three main categories

	Freq.	%
m	11553	76.50
oi	1144	7.58
ot	2405	15.93
Total	***15102***	***100***

As we have seen in Chapter 2, Sinclair and Mauranen (2006) argue that, even at the most general level of categorization, Linear Unit Grammar (LUG) analysis can provide us with a broad idea of the type of discourse under examination. For example, from the very brief samples that they provide, Sinclair and Mauranen propose that traditional published written discourse, such as that seen in a newspaper, is high in qualified m elements supported by a high number of ot elements, with an absence, or possibly very low number, of oi elements. At the other extreme, spontaneous and informal conversation is characterized as being high in oi elements and relatively low in m and ot elements, though a number of mf, ma and mr elements are also present.

Whilst it would be difficult to make any precise comparison with a comparable corpus-based study employing LUG categorization based on the figures presented in Table 5.1, if we compare the IMDb corpus discourse to the limited samples provided by Sinclair and

Mauranen, the high number of m elements, accompanied by a relatively high number of ot elements would seem to indicate that the IMDb corpus discourse is a discourse type which is, for the most part, oriented towards textual-focused relations. It would, thus, seem to confirm a commonality with 'traditional' written discourse. However, it should be noted that a substantial number of elements are situation-focused interactive oi elements, at 7.58%, indicating that it is also a discourse type with a certain amount of overt interaction. All of this would seem to corroborate previous characterizations of online message board discourse (Collot and Belmore 1996; Crystal 2001; Claridge 2007), i.e. that it can be categorized as being an interactive written discourse.

5.2.2 Relative distribution of discourse reflexive elements in m, oi and ot elements

Research Sub-question 1b: *Are DR elements distributed in the same way as non-DR elements or are there significant differences in their distribution?*

Following the definition of discourse reflexivity presented in section 4.4, 839 elements out of a total of 15,102 elements were categorized as being discourse reflexive. This represents 5.56% of all elements, as seen in Table 5.2. In order to compare the significance of this figure with other studies, it is necessary to convert this figure into a more conventional means of expressing frequency generally employed in corpus-based studies. The most typical means of expressing frequency in such studies is an occurrence of an item per 10,000 words. Hence, the figure of 839 DR elements in the IMDb corpus can be expressed as the equivalent of 204 occurrences per 10,000 words. As a point of comparison, we can note that Hyland (1998b) reports that the number of hedging devices

found in research articles was at 167 per 10,000 words. From this, Hyland concludes that hedging is central both to the genre and how the writer aligns him/herself in the academic discourse community. DR elements in the IMDb corpus are therefore more frequent than hedges in Hyland's study. The frequency of DR elements in the IMDb corpus discourse is also much higher than the frequency of the sum of all categories of metadiscourse used by Swedish writers of English, as reported by Ädel (2006). Ädel reports the overall total being around 101 occurrences of metadiscourse items per 10,000, meaning that the DR elements in this study are around twice as frequent as the occurrence of metadiscourse in Ädel's study. Although exact comparisons with these studies would not be helpful, given the large differences in the systems of analysis presented, a figure of 204 occurrences of DR elements per 10,000 words (or, in the terms normally used in this study, 5.56% of all elements) does seem to confirm that DR elements are an important feature of the IMDb corpus discourse.

Table 5.2: Frequency of DR elements

Discourse Reflexivity	**Freq.**	***%***
DR elements	839	5.56
Non-DR elements	14263	94.44
Total	**15102**	**100**

The distribution of DR elements in the three main categories can be seen in Table 5.3. DR m elements are by far the most frequent category of DR elements at over 75% just as they are in elements in general. DR oi elements are the second most frequent category at 17.64% of DR elements, unlike elements in general, where they are clearly the third most

frequent. DR ot elements are the least frequent of the three categories at a little over 7% of all DR elements.

Table 5.3: DR elements in the three main categories

Discourse Reflexivity	**Freq.**	***%***
DR m elements	630	75.09
DR oi elements	148	17.64
DR ot elements	61	7.27
Total	***839***	***100***

These findings are significant in themselves in clarifying the concept of discourse reflexivity in this study compared to that of metadiscourse in the interactive model of metadiscourse in previous studies (see Chapter 3). As is evident from the definition provided in section 4.4, the criterion for categorizing an element as discourse reflexive in this study is that there exists a reference to the discourse considered to be clear enough to be explicit. As seen in Table 5.3, following this criterion, just over 75% of DR elements form part of the propositional material, i.e. that they are m elements in the terminology of this study, whereas slightly under 25% of DR elements are non-propositional, i.e. o elements.

Table 5.4 shows the significance of the difference in the distribution of DR and non-DR elements across the three main categories. It shows the observed frequencies in the IMDb corpus in the second and third columns and the expected frequencies of the categories as calculated, using the log-likelihood ratio, in the fourth and fifth columns. The final column shows the log-likelihood ratio value (G^2). It can be seen in the table that the

observed frequency of DR elements which are categorized as being DR m elements, is very similar, at 630, to the expected frequency of DR m elements, at around 642. This similarity is reflected in the G^2 figure for m elements, which can be seen as being close to zero at -0.23. This indicates that the there is no significant difference in the distribution of DR and non-DR elements in m elements as a whole.

Table 5.4: DR and non-DR elements across the three categories

	Observed frequencies		**Expected frequencies**		**Log Likelihood[63]**
	DR	Non-DR	DR	Non-DR	G^2
m	630	10923	641.83	10911.17	-0.23
oi	148	996	63.56	1080.44	88.10**
ot	61	2344	133.61	2271.39	-51.86**

However, a difference between the observed and expected frequencies is evident in the two organizational element categories. In particular, DR elements are clearly over-represented in the category of oi elements. If they were distributed as other elements in the corpus are, we would expect there to be around 64 DR oi elements in the IMDb corpus. In fact, there are over double that number of oi elements in the corpus, at 148. With a G^2 value of 88.10, this represents an over-representation of DR elements of very strong significance. In contrast, DR ot elements are under-represented in the corpus. If DR ot elements were distributed like other elements in the corpus, we would expect there to be around 134 DR ot elements. There is actually less than half that number, at 61 DR

[63] High significance of over/under-representation: $p < 0.001$ critical value= 10.83 / -10.83 (indicated in tables by **). Weak significance of over/under-representation: $p < 0.01$ critical value= 6.63 / -6.63 (indicated in tables by *). A positive number signifies over-representation; a negative number signifies under-representation.

ot elements in the observed frequencies. This produces a similarly robust G^2 figure of -51.86, meaning that it can be concluded with confidence that that this under-representation is not simply due to random factors.

5.2.3 Summary

It has been established then that the IMDb corpus discourse is message-oriented and supported heavily by textual organizational devices. This would seem to indicate that the IMDb discourse has much in common with traditional written texts. However, there is a significant presence of interactional organizational elements suggesting that the interactional nature of online message board discourse is reflected in the linear structure of the discourse. Discourse reflexivity is salient only in oi elements, when considering the three main categories as a whole, indicating that posters would seem to refer to the present discourse most prominently in fixed and semi-fixed interactional elements. However, such conclusions are limited by the fact that they remain general at this stage. As will be seen in Chapter 6, DR elements within the three main categories are not distributed evenly across all subcategories.

5.3 Characteristics of m, oi and ot elements

5.3.1 Introduction

In this section, the lexical characteristics of the three main categories will be presented. This will be done through the presentation of WordLists by order of frequency and KeyWord lists employing the split corpus approach outlined in section 4.5.3. In this way

the general characteristics of the corpus will be described providing a general lexical description of each of the three main element types in the IMDb corpus. In addition, the non-structural discourse function of o elements will also be presented employing an adaptation of Hyland's (2005) categorization of metadiscourse. These will then be compared to DR elements in section 5.4.

5.3.2 Lexical characteristics of m elements

A WordList of the most frequent 20 words in m elements in the IMDb corpus can be seen in Table 5.5. Although, as stated previously, it is not possible at this stage to compare this to similar studies employing a LUG approach, we can note that in general terms, the list of most frequent words in m elements is of a similar nature to the most frequent word lists in other, larger more general corpora, such as the British National Corpus (BNC) (Leech et al 2001). The exception to this would seem to be the relatively high ranking of *I* in the IMDb corpus as the third most frequent word in m elements. One of the most salient differences between the written and spoken corpora of the BNC, as illustrated by Leech et al, is the prominence of *I* and *you* in the spoken corpus when compared to the written corpus. The presence of these items in the most frequent words featuring in m elements would seem to further corroborate the prominence of immediate interactional features characteristic of spoken discourse in the IMDb corpus discourse when compared to 'traditional' written discourse.

Table 5.5: Frequency WordList: m elements

N	Word	Freq.	%
1	THE	1447	3.97

2	A	962	2.65
3	I	881	2.42
4	TO	813	2.23
5	OF	696	1.91
6	IN	581	1.59
7	IS	574	1.57
8	IT	549	1.51
9	HE	536	1.47
10	WAS	486	1.33
11	YOU	399	1.09
12	THAT	364	1.00
13	#	347	0.95
14	HIS	323	0.89
15	NOT	294	0.81
16	HAVE	288	0.79
17	FOR	276	0.76
18	THIS	239	0.66
19	ON	232	0.64
20	WITH	229	0.63

Table 5.6 shows the KeyWord list of m elements when compared to all o elements. Those items with a positive keyness figure are salient in m elements, whereas those items with a negative keyness figure are salient in o elements. The majority of the KeyWords for the elements are those typically associated with most high frequency words of larger corpora such as *the*, *a*, *was*, *is* and so on. The exceptions are the words *movie*, *film* and *movies*, reminding us that the present study is based on a relatively small corpus dedicated to a narrow topic area. The fact that *I* and *you* do not feature in the KeyWord list in Table 5.6 but are high frequency words in m elements, as seen in Table 5.5, indicates that they are relatively frequent in o elements as well. As we will see in section 5.3.3, this is accounted for by their relatively high frequency in oi elements.

Table 5.6: KeyWords: m elements vs. o elements

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	THE	1447	3.9658	44	0.937	148.67	5E-16
2	A	962	2.6366	20	0.4259	126.52	1E-15
3	WAS	486	1.332	0		118.42	1E-15
4	HE	536	1.469	3	0.0639	106.51	2E-15
5	HIS	323	0.8852	1	0.0213	69.307	2E-14
6	IS	574	1.5732	14	0.2981	68.15	2E-14
7	HAVE	288	0.7893	1	0.0213	60.996	4E-14
8	ARE	224	0.6139	1	0.0213	46.901	5E-13
9	THEY	185	0.507	0		44.903	6E-13
10	THIS	239	0.655	2	0.0426	43.563	9E-13
11	IT	549	1.5046	23	0.4898	40.358	3E-12
12	MOVIE	159	0.4358	0		38.579	8E-12
13	FILM	145	0.3974	0		35.176	5E-10
14	HIM	164	0.4495	1	0.0213	31.924	1E-08
15	AN	120	0.3289	0		29.102	7E-08
16	HAD	118	0.3234	0		28.616	9E-08
17	HAS	110	0.3015	0		26.673	2E-07
18	WITH	229	0.6276	6	0.1278	25.751	4E-07
19	SHE	106	0.2905	0		25.702	4E-07
20	ABOUT	153	0.4193	2	0.0426	24.417	8E-07
21	MOVIES	100	0.2741	0		24.245	8E-07
22	PLEASE	4	0.011	9	0.1917	-24.02	1E-06
23	UNTIL	7	0.0192	11	0.2342	-25.43	5E-07
24	ALSO	35	0.0959	21	0.4472	-25.63	4E-07
25	SINCE	11	0.0301	13	0.2768	-26.04	3E-07
26	CAUSE	3		9	0.1917	-26.33	3E-07
27	TOO	40	0.1096	23	0.4898	-26.93	2E-07
28	BEFORE	23	0.063	18	0.3833	-27.55	1E-07
29	LIKE	166	0.455	53	1.1286	-28.16	1E-07
30	SORRY	7	0.0192	12	0.2555	-28.82	8E-08
31	WHEN	73	0.2001	35	0.7453	-33.73	3E-09
32	THINK	120	0.3289	56	1.1925	-52.38	1E-13
33	EITHER	6	0.0164	18	0.3833	-52.68	1E-13
34	NOW	32	0.0877	31	0.6601	-55.19	8E-14
35	OH	4	0.011	19	0.4046	-62.29	4E-14
36	THAN	38	0.1041	38	0.8092	-69.07	2E-14
37	BLAH	3		20	0.4259	-69.84	2E-14
38	FACT	5	0.0137	22	0.4685	-70.96	2E-14
39	YES	5	0.0137	25	0.5324	-82.85	7E-15
40	MAYBE	7	0.0192	28	0.5963	-88.4	5E-15
41	AS	171	0.4687	93	1.9804	-103.6	2E-15
42	YEAH	3		30	0.6388	-111.1	2E-15
43	SO	110	0.3015	86	1.8313	-132.3	8E-16
44	THEN	14	0.0384	47	1.0009	-142.2	6E-16

45	WELL	28	0.0767	61	1.299	-161.5	4E-16
46	BECAUSE	5	0.0137	101	2.1508	-401.4	2E-17
47	THAT	364	0.9976	294	6.2606	-472	1E-17
48	OR	30	0.0822	170	3.6201	-581.7	5E-18
49	BUT	10	0.0274	317	6.7504	-1309	4E-19
50	AND	69	0.1891	885	18.846	-3521	2E-20

5.3.3 Lexical characteristics of oi elements

In this section, the lexical characteristics of oi elements will be explored through the use of WordLists according to frequency and KeyWord lists. As was shown in Table 5.1, over 23 percent of all elements in the IMDb corpus are organizational elements, with 7.58% being oi elements and 15.93% ot elements. Although both oi and ot elements are termed *action-focused* elements (see Table 2.7), they are very different in character. For instance, there is a much wider range of oi elements as compared to ot elements. There are, in fact, 441 different oi elements in the IMDb corpus (discounting 81 oi elements which are proper nouns) whereas there are 203 different types of ot elements. This means that there are around half as many oi elements in the IMDb corpus as ot elements but twice the number of element types.

Table 5.7 shows the WordList for oi elements, detailing the most frequent 25 elements. The list includes several elements which have been well-documented in a variety of theoretical frameworks as was discussed in section 2.3.4. These include *I think, well*, *maybe*, *of course, now* etc. However, it should also be noted that there are a number of element types which have not featured in such studies. These are elements which have been described as being typical of computer-mediated communication (CMC) (e.g. Crystal 2001; Tagg 2012) and have already been noted as being present in online

message board discourse in particular (Lewin and Donner 2002; Claridge 2007). These include: *LOL*, *yeah*, *haha*, *:)*, *wow*, *btw* and so on. Such items draw attention again to the fact that the IMDb corpus discourse is distinct in character from the type of academic written discourse, which has normally been the focus of studies of metadiscourse or discourse reflexivity up to this point.

Table 5.7: Frequency WordList: oi elements

oi element	**Freq**	**%**
1. name	83	7.26
2. I think	54	4.72
3. well	40	3.50
4. maybe	27	2.36
5. lol ; LOL	24	2.10
6. yeah ;yea	23	2.01
7. yes	22	1.92
8. ha ; hah; haha; hehe	21	1.84
9. anyway ; anyways	18	1.57
10. of course	18	1.57
11. :)	15	1.31
12. etc ; ect	14	1.22
13. i guess	14	1.22
14. now	14	1.22
15. at least ; atleast	11	0.96
16. to me	11	0.96
17. wow	11	0.96
18. btw ; by the way	10	0.87
19. hey	10	0.87
20. man	10	0.87

21. no	10	0.87
22. perhaps	10	0.87
23. oh ; o	9	0.79
24. apparently	8	0.70
25. I mean	8	0.70
26. OK ; okay	8	0.70
27. please	8	0.70
28. sorry	8	0.70
Others	610	54.63
Total	***1144***	***100***

As indicated above, perhaps the biggest contrast between oi elements in this study compared to ot elements is the relatively heterogeneous nature of oi elements. This heterogeneity can be seen further in the range of word classes which are represented in the category of oi elements. Table 5.8 shows the categorization of oi elements according to the word class categories provided in the *Collins Cobuild Advanced Learner's English Dictionary* (2003). It can be seen that these range from the more conventional categories of adverbials with clause, such as *perhaps* and *maybe* to newer additions to the language such as abbreviations, like *LOL* and *ROTFL*, imitations of a paralinguistic action *::sigh::*, and emoticon symbols, such as ***:-)***. As well as these, there are a wide range of other oi elements which are less fixed, which do not feature as separate dictionary entries, but which may be considered to be semi-fixed multi-word expressions (such as *it's possible to say*, *it's safe to say*) or patterns (such as v + (*that*): *I admit*, *I fear* etc.)

Table 5.8: Word classes in oi elements[64]

Word Class	**Example from IMDb corpus**
Adverb	<*Name*>, I think **maybe** I agree with you!
Exclamation	**Haha** Good call, <*Name*>
Convention (fixed phrase in conversation)	Well, **You know what I mean** hahah
Phrase	**I mean** if you've seen "V2" then you should know about the basic storyline, right?
Vocative noun	Listen ,**Crazy bitch**…you have to make up you mind
Abbreviation	I think its just a joke **lol**!
Sound word	This movie sucks **::sigh::**
Proper Noun	Gee**, <*Name*>**, where have I read that before?
Emoticon symbol	The guy was just too clean **:-)**

As a consequence of this diversity, as can be seen in Table 5.9, the top 10 most frequent oi elements account for only 23.17% of the total number of oi elements. This is represented pictorially in Figure 5.1. Furthermore, the top 30 oi elements only account for 41.03% of all oi elements. This is in contrast to the top 30 ot elements, which, as we will see in section 5.3.5, account for 87.18% of all ot elements.

Table 5.9 Frequencies of Top 30 oi elements

Number	**% of total**
Top 10 oi elements	23.17
Top 20 oi elements	34.22
Top 30 oi elements	41.03

[64] Word class categories from *Collins Cobuild Advanced Learner's English Dictionary* (2003), with the exception of the category of *emoticon symbol.*

Even from a cursory look at Table 5.8, we can see that oi elements play an important role in defining the characteristics of the IMDb corpus discourse. It is in oi elements that we can get glimpses of humour, irony, playfulness and conflict, which, as will be seen in Chapter 8, is characteristic of the IMDb corpus discourse. In addition, oi elements display spontaneity and innovation with the language in the variations in spelling of oi elements. The oi element *LOL*, for example, is most typically spelled in upper case but also appears in lower case *lol* and on one occasion *LOOOOOOOOOOOOOOOOOOOOL*. Even though these features were already noted in early studies as being a characteristic of online message board discourse (e.g. Lewin and Donner 2002; Crystal 2001; Claridge 2007; Arendholz 2013), the LUG-based characterization of online message board discourse employed here allows us to have a more precise understanding of such features, namely, that they would seem to be particularly prevalent in oi elements. It is also in oi elements that those features typically associated with CMC such as emoticons and text message acronyms are found. It could be argued, then, that despite being the least frequent of the three main elements, it is in oi elements that the IMDb corpus discourse is most clearly distinct from traditional written genres.

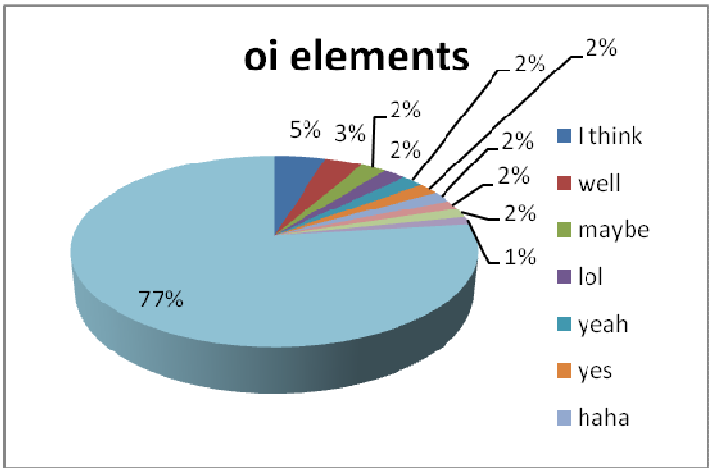


Figure 5.1: Top 10 oi elements

5.3.4 Functional characteristics of oi elements

By definition, oi elements are elements in which the pragmaticalized discourse function has precedence over the literal meaning. Thus, even though oi elements have been described in this chapter without reference to the linear structure of the discourse, mention needs to be made to the non-structural discourse function which they fulfil in order to gain a clearer picture of the type of elements they are. This could be done using a variety of theoretical frameworks. However, the most convenient to use at this juncture in this study is the categorization proposed by Hyland (2005), in particular the subcategories of interactional metadiscourse (see Table 3.3) as it offers a categorization which does not require paying attention to the structural mechanisms that will be discussed in Chapters 6–8. By employing this framework, the characteristics of these elements can be related in a non-structural way to previous metadiscourse research before we move on to examining their role in the linear structure of the discourse in Chapter 6. The categories in Table 5.10 are adapted from Hyland's and were used in characterizing oi elements in the IMDb corpus data.

Table 5.10: Non-structural discourse function of oi elements

Subcategory	Explanation
Hedges	Devices which indicate the writer's decision to recognize alternative voices and viewpoints and so withhold complete commitment to a proposition
Boosters	Words which allow writers to close down alternatives, head off conflicting views and express their certainty in what they say
Attitude markers	Devices that indicate the writer's affective rather than epistemic attitude to propositions, e.g. surprise, agreement, importance, obligation, frustration etc.
Engagement markers	Devices that explicitly address the audience either to focus their attention or include them as discourse participants
Self-mention	Occasions when the writer provides his/her own name, e.g. as signature

Table 5.11: oi elements by discourse function

	Freq.	**%**
Hedges	203	17.74
Boosters	120	10.49
Attitude Markers	346	30.24
Engagement Markers	465	40.65
Self-mention	10	0.87
Total	***1144***	***100***

It can be seen in Table 5.11 that the most common discourse function of oi elements in the IMDb discourse is as an engagement marker at over 40% of all oi elements. These include such items as the phrase *I mean*, the adverbial *well*, the interjections *Oh*, *hey* and *yeah* and so on. Clearly, these reflect the immediate interaction between participants,

which is evident in the IMDb corpus data. The second most common discourse function category for oi elements is the category of attitude markers including *Wow*, *LOL* and *Ha ha* as well as a range of emoticon symbols. Again, the interactive nature of online message board discourse is evident in these elements in that many would seem to be reactions to previous posters' contributions. The prominence of attitude markers also confirms the fact that it is a message board site whose main purpose is to provide an opportunity for posters to voice their opinions. The categories of hedges and boosters are less represented in oi elements, although hedges include the very frequent element *I think*. The prominence of *I think* as the 'prototypical expression of stance' in the IMDb corpus would seem to be consistent with other research studies into other forms of CMC, such as blogs (Myers 2010). Other hedging devices present in oi elements include *I guess*, *maybe* etc. Boosters are less frequent than hedges and include elements such as *of course* and *in fact*.

It is difficult to compare these findings with previous studies that use such a framework (e.g. Hyland 1998b, 1999, 2005; Hyland and Tse 2004) due to the difference in the categorization of the elements. Most obviously, the categories of hedges and boosters in Hyland's taxonomy include a large number of expressions of modality such as modal verbs, which in this study are, for the most part, categorized as being part of m elements. However, in general, it can be seen that, in the IMDb corpus discourse, attitude markers and engagement markers are more frequent in oi elements than hedges and boosters, whereas in Hyland's various studies of academic written discourse it is invariably the

category of hedges which is the most frequent. What proportion of DR oi elements belong to the same discourse function categories will be discussed in section 5.4.3.

5.3.5 Lexical characteristics of ot elements

As we have seen in Table 5.1, a large number of ot elements are present in the IMDb corpus. The majority of ot elements, as we can see in Table 5.12, are one or two-word elements comprising, conjunctions, sentence adverbials and prepositions. Thus, ot elements can be characterized as being a small number of reasonably homogenous elements, which are themselves dominated by a very limited number of high frequency elements, with *and*, *but* and *that* being the three most common (see Table 5.12 and Figure 5.2). In contrast, those ot elements which are characteristic of academic writing, such as formal sentential adverbials (Biber et al 1999:880) like *furthermore* and *nevertheless* play only a minor role in the IMDb corpus discourse.

Table 5.12: Top 20 most frequent ot elements

ot element	Freq.	%
1. and	849	35.30
2. but	296	12.31
3. that	249	10.35
4. or	155	6.44
5. so	74	3.08
6. because ; b/c	70	2.91
7. like	40	1.66
8. then	34	1.41
9. than	33	1.37
10. when	33	1.37
11. though ; tho	27	1.12
12. too	26	1.08
13. however	20	0.83

14. as	18	0.75
15. either	16	0.67
16. &	14	0.58
17. before	14	0.58
18. also	13	0.54
19. although	12	0.50
20. cause ; cos; cuz	12	0.50
Others	400	16.63
Total	***2405***	***100***

Thus, as can be seen in Table 5.13, the ten most frequent ot elements account for 76.03% of all ot elements. The thirty most frequent ot elements account for 87.18% of all ot elements. The dominance of a small number of extremely high frequency elements is in contrast to the diversity seen above in section 5.3.3 in oi elements.

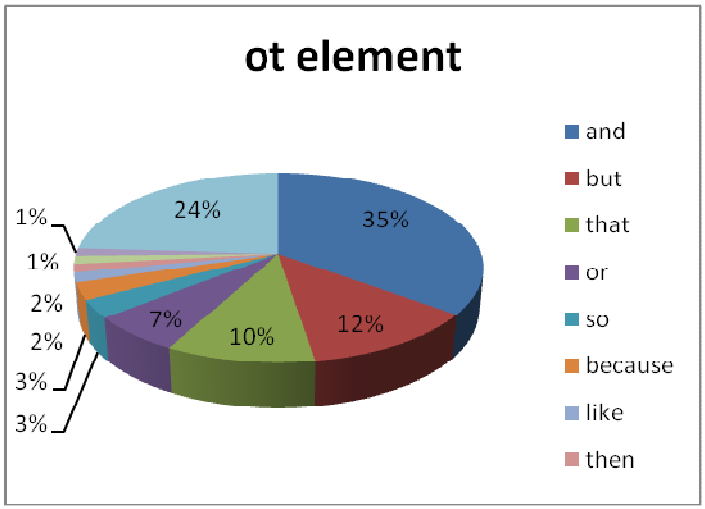


Figure 5.2: Top 10 ot elements

Table 5.13 Frequencies of Top 30 ot elements

	Freq.	**% of total**
Top 10 ot elements	1795	76.03
Top 20 ot elements	1967	82.66
Top 30 ot elements	2063	87.18

5.3.6 Functional characteristics of ot elements

As seen with oi elements, the discourse function of ot elements has precedence over any literal meaning that these items may possess. As we have seen in Chapter 3, the roles in discourse of ot elements are well established in such areas as the study of cohesion (Halliday and Hasan 1976) and in taxonomies of metadiscourse as textual metadiscourse (e.g. Vande Kopple 1985; Hyland 1998b) or interactive metadiscourse elements (Hyland 2005). An adaptation of the categories of interactive metadiscourse presented by Hyland in Table 3.3 can be used here. The three categories which appear in Table 5.14 were used in characterizing ot elements in the IMDb corpus data.

Table 5.14: Non-structural discourse function of ot elements

Subcategory	Explanation
Logical connectives[65]	Mainly conjunctions and adverbial phrases which help readers interpret pragmatic connections between elements: addition, comparison, and consequence
Frame markers	Signal text boundaries or elements of schematic text structure.
Code glosses	Supply additional information by rephrasing, explaining or elaborating what has been said to ensure the reader is able to recover the writer's intended meaning.

Table 5.15: ot elements by discourse function

	Number	Percentage
Logical Connectives	2241	93.18
Frame Markers	102	4.24

[65] The term *logical connective* is preferred here as ot elements refer to elements which link both elements and linear units whereas in Hyland (2005) the equivalent category *transition markers* refers to between propositions. Also missing from Hyland's (2005) categories are *endophoric markers* and *evidentials,* which tend to be m elements in this model.

Code Glosses	62	2.58
Total	***2405***	***100***

It can be seen from Table 5.15 that, overwhelmingly, the most frequent category of ot elements in the IMDb corpus discourse is logical connectives at over 93%. In contrast, frame markers and code glosses only account for a little over four and two percent respectively. Although the category of logical connectives or its equivalent has tended to be the most frequent as seen in studies of academic writing (e.g. Hyland 1999, 2005), it has not dominated to the extent seen in the IMDb corpus as evidenced by Table 5.15. This is partly due to differences in the theoretical frameworks of this study compared to studies of metadiscourse. The reduction from five categories in Hyland (2005) to three categories here may have had some influence. Moreover, studies of metadiscourse have tended to limit the concept to links between sentences or propositions whereas the figures in Table 5.15 include ot elements between all elements. However, combined with the information provided in Table 5.12, it can still be concluded that ot elements in the IMDb corpus discourse, in the vast majority of cases, provide a logical connection between elements and this is done in the majority of cases by the use of *and*, *but*, *that* and *or*. The increased prominence of these extremely high frequency conjunctions can therefore be considered to be a defining feature of IMDb corpus discourse.

In the next section, the lexical composition of DR elements in general and in each of the three main categories will be examined.

5.4 Characteristics of discourse reflexive m, oi and ot elements

5.4.1 Characteristics of discourse reflexive m elements

In section 5.3, a general description of the lexical characteristics of the three main element categories was presented as well as the functional characteristics of o elements. This can now be compared to the lexical characteristics of DR elements which occur in each of the three main categories and to the functional characteristics of DR o elements. Table 5.16 shows the 25 most frequent words in DR m elements. A cursory comparison between this table and the most frequent words in m elements in general in Table 5.5 reveals, as one might expect, that certain words related to communication and the production of text, such as *say*, *question*, *said* and *point* are prominent in DR m elements. However, what is also noticeable is the increased prominence of *you* in DR elements, being the second most frequent word in DR elements. In contrast, it is only the eleventh most frequent word in all m elements as seen in Table 5.5.

Table 5.16: Top 25 most frequent words in DR m elements

N	Word	Freq.	%
1	I	87	4.11
2	YOU	69	3.25
3	TO	68	3.21
4	THE	60	2.83
5	A	54	2.55
6	SAY	38	1.79
7	THIS	31	1.46
8	IT	30	1.41
9	THAT	25	1.18
10	YOUR	23	1.08
11	MY	22	1.04
12	IS	20	0.94
13	ON	19	0.90

14	IF	18	0.85
15	IN	18	0.85
16	JUST	17	0.80
17	QUESTION	17	0.80
18	SAID	17	0.80
19	ARE	15	0.71
20	OF	15	0.71
21	WAS	15	0.71
22	WHAT	15	0.71
23	CAN	14	0.66
24	HERE	14	0.66
25	POINT	14	0.66

A clearer picture of which words are most salient in DR m elements can be gleaned by drawing up a KeyWord list of DR m elements versus non-DR m elements, as seen in Table 5.17. Again, it can be seen that words related to text production and communication (e.g. *say*, *question*, *said*, *comments* etc.) feature prominently. However, it is also confirmed that certain other words not exclusively related to text production or communication are also prominent. Most obviously, the second person pronoun *you* is the third most salient word of DR m elements compared to non-DR m elements with other personal pronouns and determiners also featuring, namely *your*, *my*, *I* and *ur*. These will be discussed further below.

Table 5.17: KeyWord list: DR m elements vs. non-DR m elements

N	**Key word**	**Freq.**	**%**	**RC. Freq.**	**RC. %**	**Keyness**	**P**
1	SAY	38	1.7916	28	0.0815	130.18	9E-16
2	QUESTION	17	0.8015	2		84.308	6E-15
3	YOU	69	3.2532	333	0.9691	64.938	3E-14
4	HERE	14	0.6601	10	0.0291	48.335	3E-13
5	POINT	14	0.6601	13	0.0378	43.901	8E-13
6	SAID	17	0.8015	27	0.0786	41.363	2E-12
7	COMMENTS	7	0.33	0		39.851	4E-12
8	ANSWER	8	0.3772	1		39.388	5E-12

9	POST	11	0.5186	8	0.0233	37.733	2E-11
10	WORD	8	0.3772	2		35.778	1E-10
11	POSTS	9	0.4243	4	0.0116	35.674	2E-10
12	CALLING	7	0.33	1		33.942	3E-09
13	NAME	12	0.5658	15	0.0437	33.035	6E-09
14	YOUR	23	1.0844	78	0.227	31.984	1E-08
15	MENTIONED	7	0.33	2		30.555	3E-08
16	INSULT	6	0.2829	1		28.533	9E-08
17	MEAN	11	0.5186	18	0.0524	26.293	3E-07
18	SAYING	8	0.3772	7	0.0204	25.656	4E-07
19	MY	22	1.0372	89	0.259	25.429	5E-07
20	COMMENT	6	0.2829	2		25.397	5E-07
21	FORUM	5	0.2357	1		23.173	1E-06
22	QUOTE	5	0.2357	1		23.173	1E-06
23	I	87	4.1018	791	2.302	23.122	2E-06
24	ABOVE	4	0.1886	0		22.767	2E-06
25	STATEMENT	4	0.1886	0		22.767	2E-06
26	UR	7	0.33	7	0.0204	21.279	4E-06
27	CALL	6	0.2829	4	0.0116	21.173	4E-06
28	THREAD	6	0.2829	5	0.0146	19.595	1E-05
29	EXAMPLE	5	0.2357	3		18.234	2E-05
30	MESSAGE	4	0.1886	1		17.882	2E-05
31	INSULTS	4	0.1886	1		17.882	2E-05
32	BOARDS	4	0.1886	1		17.882	2E-05
33	GOSSIP	3	0.1414	0		17.074	4E-05
34	DISCUSSION	3	0.1414	0		17.074	4E-05
35	ATTACKING	3	0.1414	0		17.074	4E-05
36	DEBATE	3	0.1414	0		17.074	4E-05
37	THEORY	3	0.1414	0		17.074	4E-05
38	SUGGEST	3	0.1414	0		17.074	4E-05
39	THIS	31	1.4616	209	0.6082	16.867	4E-05
40	OPINIONS	5	0.2357	4	0.0116	16.573	5E-05
41	LIST	6	0.2829	8	0.0233	15.99	6E-05
42	TOPIC	4	0.1886	2		15.368	9E-05
43	OF	15	0.7072	682	1.9848	-22.54	2E-06
44	HE	5	0.2357	531	1.5453	-35.7	2E-10

The KeyWords can themselves be categorized according to the definition of discourse reflexivity as provided in section 4.4, i.e. whether it constitutes a reference to the present text; to a discourse event within the present discourse; to a discourse act within the present discourse; to the discourse participants; or to place time or manner within the present discourse. As seen in Table 5.18, the subcategory of DR m elements that makes

an explicit reference to the present text includes those which are specifically referent to CMC and/or online message board discourse. These include *post*; *posts*; *forum*; *thread*; *message boards* and *link*. Those DR m elements which include an explicit reference to the discourse event include nominalizations of verbal processes. Some of these are neutral in their reference, including: *comment*, *discussion* and *statement*. Others point to the apparent antagonism expressed through discourse reflexivity. These include: *insult*, *gossip*, *attacking, accusation, name calling* and *rant.* The noun *astroturfing*, for instance, is both antagonistic and specific to CMC, referring as it does to an activity whereby posters mimic spontaneous grassroots mobilizations in their communications but are in fact well-orchestrated campaigns mounted by companies and organizations (Monbiot 2010). It is therefore in DR discourse event nouns that we can see indications of antagonism. DR m elements containing text nouns, on the other hand, generally betray little of such conflict through discourse reflexivity. The use of discourse reflexivity as a means to express antagonism is discussed more fully in Chapter 8.

Table 5.18: DR m elements by reference type

	Examples from KeyWord list		**Further examples from IMDb corpus**	
Text	answer boards forum list message name	post posts question thread word	English board screen (name) Christmas (list)	link sentence words petition questions
Discourse event	calling comment comments debate discussion example	gossip insult insults point quote statement	accusation astroturfing delivery examples joke name calling points	pun rant reaction response rhetoric speculation

Discourse act	answer attacking call insult mean mentioned point	post question quote said say saying suggest	applaud badmouth berate blame defend discuss dissect explain flaming interpret joke name make it clear make (a good point) use (a *beep*) changing (the subject) express (their feelings) put someone down	notice point it out question ramble read reply re-read shut up talk talking threaten
Discourse participants	I my ur	you your	IMDb troll trolls	
Discourse place/ time/ manner	above here	this	there	

It would seem, then, that a broad distinction can be made between DR m elements containing discourse event nouns as being those which reflect conflict and antagonism and DR m elements containing text nouns, which are generally attitudinally neutral. This is even more pronounced on occasions when these nouns are accompanied by pre-modifying adjectives.

Table 5.19: DR nouns phrases in DR m elements

	Examples from KeyWord list		**Other examples from IMDb corpus**	
Discourse event	legitimate good plot	point	intelligent	response

	narrow minded futile oh so witty	comments	false	accusation
		insult		points
	snide ludicrous	comment	a knee-jerk visceral rejecting	reaction
	childish uneducated	statement		rant
	childish	insults	never-ending	speculation
	endless off-topic	discussion		joke
	name	calling		
	less exact Aliens Vs Predator	example	countless	examples
		message	same	rhetoric
	big	debate		
	juicy lousy	gossip		
	objective other peoples	opinions		
Text	same original simple very very hard tough good/evil his son's	question	writtin	English
		answer	message	board
	previous great original first	post		screen (name)
	strong descriptive	words	Good nice	Christmas (list)
	three previous	posts	same	sentence
	screen	name		petition
	BH IMDb religion	forum	irrelevant	questions
		quote		words
		thread		
	message	boards		
	Christmas	list		

Such nouns along with their accompanying pre-modifying adjectives or nouns can be seen in Table 5.19. Some of these pre-modifiers are clearly attitudinal expressing an

evaluation of the referent: *lousy*, *snide*, *intelligent* and *irrelevant* are examples of these. Others are non-attitudinal such as *previous*, *Christmas* or *writtin* [sic]. Most nouns in the table which have an attitudinal pre-modifying noun are discourse event nouns and most of these provide a negative evaluation of the discourse event. As Hoey (2000) argues, the placing of the adjective in the premodifying position indicates to the reader that the evaluation is less negotiable than if it were in the position of complement after a link verb. Thus, the evaluation of the discourse event through the ADJ+N structure provides the writer with a strong evaluative unit. The tendency to express negative evaluation in an ADJ+N pattern would seem to be stronger in some nouns than in others. In Table 5.19 those words which are shaded indicate a negative evaluation of the text or discourse event. It would seem, for example, that the noun *comment/comments* has a tendency to be proceeded by a negative adjective. The examples in the IMDb corpus include *narrow minded comments, snide comment*, *ludicrous comment* and the ironic *oh so witty comments*. Other negative premodifying adjectives include those which are related to a negative evaluation of another posters' contribution in terms of it not meeting standards of acceptable behaviour in the discourse. These relate to the contribution not being considered, calm or rational such as *a knee-jerk visceral rejecting reaction* and to not being mature such as *childish uneducated statement* and *childish insults*. This again is an indication of the preponderance of DR elements at points of antagonism and conflict in the discourse. How this relates to the linear structure of the discourse will be discussed in Chapter 8.

Returning to Table 5.18, it can be seen that discourse act words which feature in DR m elements can be categorized in a similar fashion to elements containing text and discourse event nouns. The most clearly discourse reflexive verbs are those associated with communication such as: *say*, *mean*, *mention*, *answer* etc. These are what Ädel (2006) terms *verba dicendi* by which the fact that something is being communicated is highlighted. However, in the IMDb corpus data, the discourse act verbs present go beyond the high frequency rather neutral verbs presented by Ädel. In particular, discourse act words in the IMDb corpus also include verbs specific to CMC such as *post* and *flame*. The discourse act of flaming, i.e. 'displaying hostility by insulting, swearing or using otherwise offensive language' (Moor et al 2010) is one of a number of discourse act verbs in DR m elements which again display conflict and antagonism. As well as *flame*, the IMDb corpus includes: *attack*, *insult*, *badmouth*, *belittle*, *berate*, *blame*, *ramble, say sh*t, threaten,* and *use a *beep**. All of these indicate points of conflict in the text where the poster is characterizing a previous contribution by another poster in a negative light.

The fact that the pronouns, *you*, *I*, *my*, *your* and *ur* all feature as KeyWords for DR m elements in the IMDb corpus data would also seem to support Ädel's categorization of metadiscourse with its emphasis on personal metadiscourse and the occurrences of *I* and *you*. It is therefore important to consider for which verbs in Table 5.17 these pronouns act as subjects. Table 5.18 shows the DR verbs which feature in the KeyWord list in Table 5.17 according to the subject of the verb. It can be seen that this varies from verb to verb. The most common verb, *say*, for example, would seem to more commonly associated with the subject *I* in m elements. As will be seen in 5.4.2, this may differ from oi

elements containing *say*. Other verbs, such as *mean*, *said* and *call* would seem to be associated more closely with the subject *you,* at least according to the limited evidence available here. This is explored further in section 6.6.2. Again, the oi elements containing these verbs may have a different association. Most obviously, *mean* combines with the subject *I* in the frequent oi element *I mean* (see section 5.4.2).

Table 5.20: DR verbs according to subject

	I	**You**	**We**	**Others**	***Total***
say	24	6	1	2	***33***
mean	3	7		1	***11***
saying	2	3		3	***8***
mentioned	2	2		2	***6***
said	1	4			***5***
answer	1	3			***4***
call		4			***4***
insult	1	2			***3***
attacking	1			2	***3***
suggest	1	2			***3***
point		1			***1***
post	1				***1***
quote				1	***1***
question		1			*1*
Total	***37***	***35***	***1***	***11***	***84***

It is also of note that DR m elements in the IMDb corpus contain a number of references to discourse participants. As established in Chapter 1, IMDb discourse is potentially polylogic in nature, i.e. that more than two posters are contributing often with more than one poster replying to one turn. This means that the roles of the different posters can be quite complex. A poster can simply be responding to a previous poster's contribution

such that the previous poster is the addressee. In Example 5.1 the poster is responding directly to a previous poster's post:

Example 5.1[66] [67]

P	T	L	Element	Element type
A	1	1.	The Worst 'Best' Oscar Movie Ever??	|m
B	2	2.	**How can you say**	**|m**
		3.	this is	m–
		4.	the worst Oscar movie ever?	+m

In this example the addressee, *you* in Line 2, clearly refers to the Poster A. However, there are also occasions in the IMDb corpus discourse where the poster responds to the previous post by commenting about the post to other posters or potential posters.

Example 5.2[68]

P	L	Element	Element type
A	1	**LOL-**	**|oi**
	2	I think	|oi
	3	**the OP needs glasses.**	**m**

In Example 5.2 the *OP* (the *Original Poster*) in Line 3 is relegated to being an 'overhearer' in Goffman's terms (Levinson 1983). In layman's terms, the previous poster, the OP in this case, becomes what is being talked about rather than being talked to.

[66] As seen in Chapter 4, in these element example tables, double horizontal line indicates the end of a turn; a solid horizontal line indicates the end of a subject line; a dotted line indicates the end of an element. A wavy line means that part of the same turn is omitted. A double wavy line means that part of discourse is missing from another turn. The posts are divided into elements, one per line.
[67] This example is from Thread 067 about the film *Braveheart*.
[68] This example is from Thread 039 about American actress Uma Thurman in which the original poster posts a link to a photo of himself asking if other posters think he looks like Uma Thurman.

However, the overhearer can just as easily then respond directly to the post about him/herself which was directed to the other posters. This will be discussed in Chapter 8.

There are also several terms referring to the discourse participant which reflects the conflict and antagonism in the IMDb discourse which has already been referred to at various points in this chapter already. Most noteworthy is the labelling of some discourse participants as being a *troll*. This is a CMC-specific DR term, referring to a poster who is guilty of committing personal attacks or behaving in a manner likely to provoke a negative response (Arendholz 2013:48). It is a subject judged to be of great importance for IMDb evidenced by the fact that, as administrators of the message boards, the website provides an extended definition of the term *troll* in the *Terms and Conditions* for users of the message boards with a series of warnings for misconduct. Despite there being some academic debate to the extent that any poster who is seemingly aggressive or provocative can be characterized as being a *troll* (e.g. Moor et al 2010), it is clear that the term acts as a provocative label and is well established as such among the posters on the IMDb message boards.

Thus, we have seen that discourse reflexivity manifests itself in a variety of ways in m elements in the IMDb corpus discourse. This includes labelling and evaluating a discourse event or part of the text. Evidence has also been presented in this section that discourse reflexivity in m elements is often associated with both the specifics of online message board communication and also antagonism and conflict. This will be explored more fully in Chapter 8.

In the next section, the lexical composition of oi elements will be examined.

5.4.2 Lexical characteristics of discourse reflexive oi elements

In this section, the lexical characteristics of DR oi elements will be presented. This will be done following the same approach as employed in section 5.3.3. Firstly, the lexical properties of DR oi elements will be presented through the use of WordLists according to frequency and KeyWord lists. In addition, in section 5.4.3, the discourse functions of DR oi elements will be presented.

Table 5.21: Top 10 most frequent DR oi elements

oi element	**Freq.**
lol / LOL / LOOOOOOOOOOOOOOOOOOOL	23
I guess / i guess / Guess	16
i mean/ I mean	9
wait / ok wait / no wait	6
Again / Once again / Again?	4
Honestly	3
QUOTE	3
Well said	3
I admit	2
I must admit	2
I SWEAR / I swear	2
Just kidding	2
lets say /Let's say	2
ROFL	2
seriously	2
you know what I mean /u know what i mean	2

First of all, then, the lexical properties of DR oi elements will be considered. Table 5.21 shows the most frequent DR oi elements in the IMDb corpus discourse. As can be seen, the most frequent DR oi element is *LOL*. This is an acronym meaning 'Laughing Out Loud' (Tagliamonte and Denis 2008) and is frequent in instant messaging and other forms of CMC. In terms of its discourse reflexivity, the verb *laugh* in this context is conceived as being a discourse act verb of communication reacting to the present discourse. Other similar oi elements present in the data are *ROTFL* (Roll On The Floor Laughing); *LMAO* (Laughing My Ass Off) and the compound acronym *ROTFLMAO* (Roll On The Floor Laughing My Ass Off).

Apart from DR oi elements made up of such acronyms, almost all other most frequent DR oi elements contain a discourse act verb. Most commonly, these are combined with the first person pronoun as subject. As we have seen in DR m elements, these tend to be *verba dicendi* such as *guess*, *mean*, *said*, *admit*, *swear* and *say*. Some of these DR oi elements can be considered to be fixed phrases such as *I guess* and *I mean* and *you know what I mean*. Other discourse act verbs in the table do not combine with the first person singular. The oi element *wait*, for instance, is an imperative normally related to trying to control the actions of the other discourse participants in the real-time nature of spoken interaction by controlling the turn taking. As commented by Tagg (2012), such interactive elements are frequently used in written CMC discourse for rhetorical effect rather than being related to such functions normally associated with real-time spoken interaction.

Table 5.22: DR oi elements by reference type

	Most Frequent oi elements	
Text	~	
Discourse event	QUOTE	
Discourse act	LOL I guess I mean wait Well said I admit	I must admit I swear Just kidding Let's say ROFL you know what I mean
Discourse participants	**I** guess **I** mean **I** admit **I** must admit **I** swear	Let**'s** say you know what **I** mean
Discourse place/ time /manner	Again Honestly seriously	

Other than discourse acts, the other type of DR oi element present in the most frequent elements are those which make reference to the manner of communicating. These include *again*, which refers to the fact that the upcoming or previous element sequence is similar to something stated previously in the present discourse. This category also includes certain style stance adverbials, namely *honestly* and *seriously* which 'convey a speaker's comment on the style or form of the utterance, often clarifying how the speaker is speaking or how the utterance should be understood' (Biber et al 1999:764).

Table 5.23 shows the KeyWord list of DR oi elements versus non-DR oi elements. For the most part, the list confirms the salience of the discourse verbs discussed above, with the addition of *tell* as well as the salience of the first person singular pronoun *I* in DR oi elements. Other examples of DR oi elements containing the first person singular includes: *I can assure you*; *I get the impression*; *I hate to be the one to break it to you*; *I know what*

you mean; *I say*; *I should point out*; *I won't lie*; *i would assure you*; *if i can call them that*; *if u know what im saying*; *i'm paraphrasing here*; and *Maybe I'm showing my age here*.

To sum up, for the most part DR oi elements in the IMDb corpus are predominantly multi-word elements containing either a verb and/or a subject pronoun, most frequently the first person pronoun. These are supplemented by the presence of less frequent one-word style stance adverbials and CMC acronyms.

Table 5.23: KeyWord list: DR oi elements vs. non-DR oi elements

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	LOL	22	6.044	0		74.595	1E-14
2	GUESS	16	4.3956	0		54.027	1E-13
3	MEAN	14	3.8462	0		47.209	3E-13
4	SAY	12	3.2967	0		40.41	3E-12
5	BLAH	16	4.3956	4	0.2548	35.615	2E-10
6	WAIT	6	1.6484	0		20.123	7E-06
7	TELL	5	1.3736	0		16.758	4E-05
8	I	42	11.538	83	5.2866	16.575	5E-05
9	AGAIN	4	1.0989	0		13.397	0.0003
10	ADMIT	4	1.0989	0		13.397	0.0003
11	SAID	4	1.0989	0		13.397	0.0003
12	HERE	4	1.0989	0		13.397	0.0003

5.4.3 Functional roles of discourse reflexive oi elements

The second approach to the description of DR oi elements is to consider their discourse function employing the same subcategories as was seen in section 5.3.4. Table 5.24 shows the significance of the difference between the frequency of DR oi elements and non-DR oi elements in the corpus according to the five discourse function categories. It can be seen that DR oi elements are over-represented significantly in three subcategories:

attitude markers; engagement markers; and hedges. Indeed, it is the category of attitude markers which is the most frequent DR oi element category whereas for non-DR oi elements it is clearly the category of engagement markers which is the most frequent.

Table 5.24: DR and non-DR oi elements by discourse function

	Observed frequencies		**Expected frequencies**		**Log likelihood**
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
hedge oi	27	176	11.28	191.72	17.02**
booster oi	13	107	6.67	113.33	5.06
attitude marker oi	63	283	19.22	326.78	68.16**
engagement marker oi	45	420	25.83	439.17	12.47**
self-mention	0	10	0.56	9.44	-1.14

The most frequent attitude makers, as seen in Table 5.25, include the CMC acronyms discussed above, most prominently, *LOL*, *ROTFL* etc. They also include certain fixed or semi-fixed phrases, such as *well said* and *just kidding* and those sentence adverbs described above: *again*, *honestly*, *seriously*. DR oi elements acting as engagement markers are also prominent in the IMDb corpus data with the most frequent element being the fixed phrase *I mean.* As we have seen, DR oi hedges are less frequent but are still significantly over-represented in the IMDb corpus. By far the most frequent DR oi element is *I guess*. These will be discussed in relation to the oi element subcategory which they represent and their role both within the linear unit and between linear units in Chapter 6.

Table 5.25: DR oi elements by discourse function

hedges oi	**Freq**	**boosters oi**	**Freq**	**attitude markers oi**	**Freq**	**engagement markers oi**	**Freq**

I guess	15	I swear	2	LOL	23	I mean	9
QUOTE	3			I (must) admit	4	wait	6
				ROTFL	4	Let's say	2
				honestly	3	You know what I mean	2
				well said	3		
				again	3		
				just kidding	2		
				seriously	2		

We may also consider DR oi elements by examining the elements which include the most salient words according to Table 5.23 and how they relate to Hyland's discourse functional categories. It can be seen in Table 5.26 that DR oi elements which contain certain DR words belong to the same discourse function. Examples of these are DR oi elements containing *guess* (*I guess* and *guess*), which would all seem to act as hedges; DR oi elements containing *admit* (*I admit* and *I must admit*) would similarly all seem to be attitude markers as would *said* (*well said*, *Couldn't have said it better myself*).

Table 5.26: DR oi elements containing KeyWords by discourse function

DR word	Example	Freq.	Discourse Function
say	it's possible to say	1	Hedge
	some might say	1	Hedge
	to say the least	1	Hedge
	it's safe to say	1	Booster
	I say	1	Booster
	True say	1	Attitude Marker
	What do I say	1	Attitude Marker
	It kills me to say it	1	Attitude Marker
	As I often say	1	Attitude Marker
	as you say	1	Engagement Marker
	Let's say / lets say	2	Engagement Marker
tell	tell me / Just tell me	2	Engagement Marker

	To tell you the truth	1	Engagement Marker
	You tell them	1	Engagement Marker
	can ya tell me?	1	Engagement Marker
mean	i mean / I mean	9	Engagement Marker
	I know what you mean	1	Engagement Marker
	u/ you know what i / I mean	2	Engagement Marker
	What do u mean	1	Engagement Marker
	if that's what you mean	1	Engagement Marker
guess	I guess / i guess	13	Hedge
	Guess	2	Hedge
admit	I admit / I admit	2	Attitude Marker
	I must admit	2	Attitude Marker
said	Well said	3	Attitude Marker
	Couldn't have said it better	1	Attitude Marker
wait	(Ok/ no) wait	5	Engagement Marker
point	I should point out	1	Booster
	Great point	1	Attitude Marker
	back to the point	1	Engagement Marker
quote	QUOTE	3	Hedge
question	The big question is still	1	Engagement Marker
	A question to you	1	Engagement Marker
	To answer your question	1	Engagement Marker
LOL	LOL / lol	10	Attitude Marker
Again	Again	2	Attitude Marker
	Once again	1	Attitude Marker
	Again?	1	Engagement Marker
Here	I'm paraphrasing here	1	Hedge
	Maybe I'm showing my age here	1	Hedge
	You're sitting here telling me	1	Attitude Marker
	Someone correct me here	1	Engagement Marker

Although the DR oi elements containing *admit* and *said* are attitude markers, the difference between the two is also clear. DR oi elements including *admit* provides an attitudinal preface for the poster's own upcoming series of elements as seen in Example 5.3.

Example 5.3[69]

P	L	Element	Element type
A	1.	**I must admit**	**|oi**
	2.	my sister does think	|m
	3	i'm strange	m
	4.	for liking the classics	ms

On the other hand, the DR oi elements above containing *said* provide a retrospective evaluation of a previous poster's post as seen in Example 5.4. This is directly related to the subcategory of the oi element involved, its position within the linear unit and the position of the linear unit in the turn. This will be discussed more fully in Chapter 6.

Example 5.4[70]

P	**T**	**L**	**Element**	**Element type**
A	1	1.	I'm gay	m
		2.	(and	+ot–
		3	Australian	ms
		4.	not that it matters)	ois
		5	and	+ot–

[69] This example is from Thread 010 about American actor Humphrey Bogart and people's attitudes to young people nowadays liking the actor.
[70] This example is from Thread 058 about the film *Borat* and its supposed homoerotic content.

		6.	loved the film	ms
B	2	7.	**Well said,**	**oi**
		8	<name>	ois

Nevertheless, it can also be seen in Table 5.26 that different DR oi elements containing the same DR word can fulfil a range of non-structural discourse functions. For example, DR oi elements which contain the verb *say* include hedges such as *some might say* and boosters such as *I say* (as seen in Example 5.5).

Example 5.5[71]

P	Line No	Element	Element type
A	1.	**I say**	**loi**
	2.	quit trying	m–
	3	to stop time	+m
	4.	and	+ot–
	5	just	loi
	6.	get on with it	m

DR oi elements containing *say* also include attitude markers such as *True Say*, which is similar to *Well said* in Example 5.4 in that it is reacting to another poster's contribution. They also include engagement markers like *Let's say* as seen in Example 5.6 where *Let's say* is employed by the poster as a means of seeking agreement with the other posters about the delimitation of the topic for discussion.

Example 5.6[72]

[71] This example is from Thread 027 about American actor Brad Pitt and whether he is aging badly.

P	Line No	Element	Element type
A	1.	**Let's say**	**loi**
	2.	we discount everything	m
	3	before 'Paths of Glory'	ms

To sum up, when considering the non-structural discourse functions of DR oi elements, it would seem that DR oi elements are more frequently engagement and attitude markers than hedges and boosters. DR oi elements containing certain DR words are tightly related to one discourse function, whereas others represent a variety of different discourse functions.

It has also been seen that such a non-structural description of discourse can only take us so far. It is therefore necessary to re-examine these items taking into consideration their place in the linear unit and in inter-linear unit relations. This will be done in Chapter 6.

5.4.4 Lexical characteristics of discourse reflexive ot elements

In this section, the characteristics of DR ot elements will be presented. As has been established, DR ot elements are extremely low frequency. Nonetheless, the same approach was used with these elements as was employed in section 5.4.2 with oi elements. Firstly, the lexical properties of DR ot elements will be presented through the use of WordLists according to frequency and KeyWord lists. Secondly in section 5.4.5, the

[72] This example is from Thread 058 about the American director Stanley Kubrick in which posters are asked what they think his weakest film is.

discourse functions of DR ot elements will be presented employing the same categories, namely connectives, frame markers and code glosses as seen in section 5.3.6.

Table 5.27: Top 10 most frequent DR ot elements

ot element	Freq.
1	7
2	7
for example	6
Second / Secondly/ second of all	5
3	4
compared to /kompared to/ as compared to	4
not to mention	4
First / first of all / First off	3
4	2
for instance	2
P.S	2

Table 5.27 shows the most frequent DR ot elements in the IMDb corpus discourse. As can be seen, the most frequent types of DR ot elements are cardinal and ordinal numbers. Apart from these, other frequent DR ot elements are fixed phrases such as *not to mention*, *for example* and *for instance*. This finding is consistent with previous studies such as Ifantidou (2005), who indicates that only a very few connectives (e.g. *in other words*, *for example* and *in short*) are textually explicit.

Table 5.28: DR ot elements by reference type

	Most frequent ot elements
Text	
Discourse event	for example for instance
Discourse act	not to mention (as) compared to
Discourse participants	

Discourse place/ time /manner	1 /2/ 3/ 4 second / secondly/ second of all first / firstly / first of all

The most frequent type of DR reference is that of discourse time, primarily through cardinal and ordinal numbers. These ot elements place the linear units and elements in a sort of virtual chronology within the discourse. Other than these, there are examples of discourse events references, e.g. *for example* and *for instance* and to the discourse acts of *mention* and *compare*.

Table 5.29 shows the KeyWord list of DR ot elements versus non-DR ot elements. The list confirms the salience of the words contained in the most frequent elements listed above.

Table 5.29: KeyWord list: DR ot elements vs. non-DR ot elements

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	#	20	15.267	0		125.02	1E-15
2	EXAMPLE	6	4.5802	0		36.848	3E-11
3	FOR	9	6.8702	10	0.3801	30.098	4E-08
4	TO	10	7.6336	17	0.6461	27.657	1E-07
5	MENTION	4	3.0534	0		24.506	7E-07
6	SECOND	4	3.0534	0		24.506	7E-07
7	COMPARED	4	3.0534	0		24.506	7E-07
8	INSTANCE	3	2.2901	0		18.357	2E-05
9	FIRST	3	2.2901	1	0.038	13.953	0.0002
10	NOT	4	3.0534	5	0.19	12.612	0.0004
11	AND	2	1.5267	875	33.257	-85.18	6E-15

5.4.5 Functional roles of discourse reflexive ot elements

Table 5.30 shows the significance of difference between the frequency of DR ot elements and non-DR ot elements in the corpus according to the three discourse function categories. It can be seen that DR ot elements are under-represented significantly in two of the three subcategories. DR ot elements are actually over-represented in code glosses, albeit with only weak significance, although the frequency here is very low. It is clear that DR ot elements are infrequent in all non-structural discourse functional categories. This would seem to confirm the peripheral role that ot elements play in any description of discourse reflexivity in the IMDb corpus discourse. It is of note, however, that of the low number of DR ot elements that do exist, over half of them are frame markers, whereas in non-DR ot elements frame markers represent a little over 2% of all ot elements.

Table 5.30: DR and non-DR oi elements by discourse function

	Observed frequencies		**Expected frequencies**		**Log likelihood**
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
connective ot	15	2226	124.50	2116.50	-161.08**
frame marker ot	36	66	5.67	96.33	-83.21**
code gloss ot	10	52	3.44	58.56	8.97*

This is reflected in Table 5.31 which shows the most frequent DR ot element related to the discourse functions present in section 5.3.6.

Table 5.31: Most frequent DR ot elements by discourse function

Connectives	**Freq**	**Frame markers**	**Freq**	**Code glosses**	**Freq**
compared to	4	1	7	for example	6
not to mention	4	2	7	for instance	2

		secondly	5		
		3	4		
		first	3		
		4	2		
		P.S.	2		

As can be seen, frame markers in the form of cardinal and ordinal numbers are particularly prominent. This relatively high number cardinal and ordinal numbers is a characteristic which the IMDb corpus discourse shares with academic written discourse (Biber et al 1999). In this respect, DR numerals are used in threads which have a more academic tone in the IMDb discourse. This can be seen in the extended example provided in Example 5.7, in which a reasonably well-elaborated argument is organized through the use of DR cardinal numbers.

Example 5.7[73]

With all due respect, I think you're completely off-base. IMO, here's why:
1. Ricky's "...sad, sad.." comment meant that, as you said, he was wrong but also that his father's pathetic because he'd physically harm his own son due to such an intense intolerance based on insecurity.
2. If Ricky's father was really "testing" Lester, he wouldn't have just killed him anyway. His character "investigated" and mulled over things all throughout the movie. He would've done so even more had Lester "passed the test" and turned out to be straight.
3. The entire situation brought Ricky's father's homosexual feelings to the surface. The gay neighbors and then, especially, what he thought he saw and his feelings toward his son's supposed sexuality brought his feelings to the surface and forced him to recognize them. They all just boiled over when he kissed Lester. The rejection is what made him kill Lester.
4. You can't disregard the irony of a charcter like that being gay. This film is all about American life and the human condition in general and irony is a huge part of life.

[73] This example is from Thread 062 about the film *American Beauty* and the sexual orientation of one of its main characters.

However, the use of DR numbers also reflects the non-academic nature of other parts of the IMDb corpus discourse. For example, the use of lists is frequent in the data often produced by ot elements featuring cardinal numbers. Rather than reflecting an academic style, this use of lists is more akin to the use of lists in popular culture as popularized by such pop literature writers as Nick Hornby, for instance in the novel *High Fidelity* (Hornby 1995). An example from the IMDb corpus can be seen in Example 5.8.

Example 5.8[74]

P	T	L	Element	Element type
A	1	1	My Top Five List	m
		2	**5.**	**+ot–**
		3	Good Will Hunting	m
		4	**4.**	**+ot–**
		5	Crash	m
		6	**3.**	**+ot–**
		7	The Shawshank Redemption	m
		8	**2.**	**+ot–**
		9	Braveheart	m
		10	**1.**	**+ot–**
		11	Amélie	m

To sum up, DR ot elements are generally of low frequency and are not considered to be salient in the IMDb corpus discourse. Those that are in evidence do not fulfil the discourse function of connectives, as the vast majority of ot elements do. Instead, they tend to be numerical frame markers, placing the discourse most typically in a chronological sequence such as a list.

[74] This example is from Thread 057 about American director Gore Verbinski, although the thread had gone a little off-topic by this stage.

5.5 Chapter summary

This chapter has examined the first research question of the study, i.e. the general distribution of elements in the IMDb corpus and DR elements in particular. It was found that the IMDb discourse is predominantly message-oriented as around three quarters of all elements are m elements. These are supported by a large number of ot elements. It has been argued that despite oi elements being the least frequent of the three categories, it is the category of oi elements which gives the IMDb corpus discourse its unique character.

It was found that 5.56% of all elements are DR elements in the IMDb corpus. This would seem to indicate that discourse reflexivity is an important feature. As regards their distribution, it was found that, just as with non-DR elements, around three quarters of DR elements are m elements. However, they are over-represented in the category of oi elements and under-represented in the category of ot elements.

As regards the characteristics of these elements, m elements would seem to be similar in nature to those seen in other corpora. However the character of oi elements would seem to be more specific to online message board discourse with a large number of CMC-specific elements and interactive devices, especially attitudinal and engagement markers. In contrast, ot elements are rather mundane very frequent logical connectives for the most part.

DR m elements may be characterized as comprising a large number of elements which make reference to the present text. Those that actually refer to part of the text tend to be

neutral whereas those that refer to the discourse event tend to express a negative evaluation. There are also a large number of discourse act verbs. As well as the *verba dicendi* one would expect here, there are a number of verbs that indicate the same negative evaluation. It is also true that, when we look at the category of discourse person, the complexities of polylogic discourse becomes apparent with *you* for instance sometimes being the addressee and sometimes the overhearer.

DR oi elements are characterized mostly by the prevalence of discourse verbs whether traditionally expressed or in internet acronyms. DR oi engagement markers and attitude markers are especially over-represented in comparison to their non-DR counterparts. It is also shown that a more structural description is required to examine the role of these elements as, for example, attitude markers can both act as a preface or retrospectively evaluate.

Finally, ot elements are infrequent so it is difficult to draw any firm conclusions other than to note that most examples of DR ot elements are marginal in that they do not tend to make reference to the discourse participants. However, from the limited data it can be seen that they do have an effect for example in creating the tone of the writing.

In the following chapter, the findings relating to the subcategories of elements will be presented and discussed.

CHAPTER 6

DISCOURSE REFLEXIVITY IN SUBCATEGORIES OF ELEMENTS IN THE IMDB CORPUS DISCOURSE

6.1 Overview

This chapter is divided into two main parts. Sections 6.2 to 6.4 are dedicated specifically to answering research questions 2–4 as outlined in section 4.5 through quantitative means. The second part of the chapter, section 6.5, is dedicated to a more detailed description of the most salient discourse reflexive (DR) elements as per the results from sections 6.2 to 6.4.

The research questions to be answered in this chapter, then, are as follows:

Research Question 2: *How are m elements in general – and DR m elements in particular – distributed in the IMDb corpus discourse?*

Research Question 3: *How are oi elements in general – and DR oi elements in particular – distributed in the IMDb corpus discourse?*

Research Question 4: *How are ot elements in general – and DR ot elements in particular – distributed in the IMDb corpus discourse?*

These research questions will each be divided into two research sub-questions, one relating to the elements in general from that category (i.e. both DR and non-DR elements together) and the second related specifically to DR elements from the same category. In order to answer Research Question 2 the frequency of the subcategories of m elements

will be presented in section 6.2.1. Having established the relative frequency of the subcategories of m elements in general, the distribution of the DR and non-DR elements in each m subcategory will be presented in section 6.2.2. A similar process will be carried out with the subcategories of oi and ot elements to answer Research Questions 3 and 4 in 6.3 and 6.4 respectively. The subcategories of each main element category are as follows:

1. Subcategories of m elements: core m; compliant (m–, +m–, +m, ms, ms–); and suspensive (|m, |m– and |mf) elements.
2. Subcategories of oi elements: core oi; compliant (ois); and suspensive (|oi) elements.
3. Subcategories of ot elements: compliant (+ot– and ots) and suspensive (|ot) elements.

Having established in which subcategories DR elements are most salient, these will be examined in detail in terms of their lexical characteristics. This will be done, as has been seen in Chapter 5, through both WordLists and KeyWord analysis.

The most salient words in the DR elements of each of these subcategories will be examined in terms of their meaning and their role in linear unit relations. This will be done in Section 6.5.

6.2 Subcategories of m elements

6.2.1 Subcategories of m elements in general

Research Question 2: *How are m elements in general – and m DR elements in particular – distributed in the IMDb corpus discourse?*

Research Sub-question 2a*: What are the relative frequencies of core m elements, the compliant subcategories of m elements and suspensive subcategories of m elements?*

It has been argued, based on the coarse grain analysis presented in Chapter 5, that the IMDb corpus discourse is characterized as being dominated by m elements, accounting, as they do, for 76.5% of all elements. This rather general statement can now be made more precise by examining the relative frequencies of the subcategories. These can be separated into three categories according to linear compliance: core m elements (core m); compliant m elements (m–, +m, +m–, ms and ms–); and suspensive m elements (|m, |m– and |mf) (see Table 4.1).

It can be seen in Table 6.1 that 17.40% of all m elements are core m elements; 72.91% of m elements are compliant; and 9.69% of all m elements are suspensive m elements.

Table 6.1: m elements by linear compliance

	Core	**% of m elements**
core m	2010	17.40
Total	***2010***	***17.40***
	Compliant	**% of m elements**
m–	2008	17.38
+m	2578	22.31
+m–	901	7.8

ms	2476	21.43
ms–	461	3.99
Total	***8424***	***72.91***
	Suspensive	**% of m elements**
|m	805	6.97
|m–	307	2.66
|mf	7	0.06
Total	***1119***	***9.69***

It is of note that the number of compliant m– elements, at 2,008, and core m elements, at 2,010 is almost identical. As explained in section 4.3.3.2, compliant m– element can be envisaged as representing the first element in a cleft-core element sequence. The figures indicate that the IMDb corpus discourse contains as many lengthier cleft-core elements as it does simple core elements. There is also a high number of supplementary elements (ms and ms– elements), at over 2,900. Although it is difficult to reach any conclusions about these frequencies without having another similar corpus-based Linear Unit Grammar (LUG) study with which these results can be compared, it would seem that the IMDb corpus discourse can be characterized as containing a high number of M linear units which contain lengthy cores and extensions.

It can also be seen in Table 6.1 that there is an important number of suspensive m elements present in the data. Taking all suspensive subcategories together, 9.69% of all m elements are suspensive, with |m elements being by far the more frequent of the

suspensive subcategories. These can be further divided according to their place in the linear unit, i.e. as linear unit-initial or medial suspensive elements (see Table 6.2).[75]

Table 6.2: Suspensive m elements

	Freq.	**% of suspensive m elements**
|m (initial)	685	61.22
|m– (initial)	278	24.84
|m (medial)	120	10.72
|m– (medial)	30	2.59
|mf	7	0.63
Total	***1119***	***100***

As we can see in Table 6.2, the clear majority of suspensive m elements, at over 86%, are actually linear unit-initial elements. As will be remembered from section 4.3.2.5, these are elements which begin the linear unit but which the reader knows will not in themselves reach the target state, for example, elements which include the verb in a V *that* pattern (Francis et al 1996) such as *I think* and *you say* and non-finite dependent clauses. In contrast, only a minority of suspensive elements, at around 13%, are linear unit-medial suspensive elements. Linear unit-medial suspensive m elements include structures such as non-defining relative clauses and prepositional phrases.

In the next section the distribution of DR m elements will be explored.

[75] |mf elements were not further subcategorized as their frequency is negligible.

6.2.2 Subcategories of discourse reflexive m elements

Research Sub-question 2b: *Are DR m elements distributed in the same way as non-DR m elements or are there significant differences in their distribution?*

It was already established in section 5.2.2 that there is no over-representation or under-representation of DR elements in m elements in general. However, as will be seen in this section, that conclusion does not hold for all subcategories. In fact, there exists significant variation (see Table 6.3).

Table 6.3: DR and non-DR m elements

Linear compliance	**Subcategory**	**Observed frequencies**		**Expected frequencies**		
		DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Core	**m**	146	1864	111.67	1898.33	10.24*
Compliant	**m–**	77	1931	111.56	1896.44	-12.65**
	+m	146	2432	143.22	2434.78	0.06
	+m–	42	859	50.06	850.94	-1.45
	ms	80	2396	137.56	2338.44	-29.80**
	ms–	23	438	25.61	435.39	-0.29
Suspensive	**|m**	87	718	44.72	760.28	33.63**
	|m–	29	278	17.06	289.94	7.40*
	|mf	0	7	0.39	6.61	0.80

In compliant m elements, there is a significant under-representation of DR elements in m– and ms elements with a G^2 figure of -12.65 and -29.80 respectively. Thus, DR

elements are far less frequent than would be expected if they were distributed in the same way as non-DR elements in the initial element of a cleft-core element sequence and also in self-contained supplementary (ms) elements. In the category of ms elements, for instance, it can be seen in Table 6.3, that the expected frequency of DR ms elements is around 138, whereas the observed frequency of DR ms elements is only 80. In the other subcategories of compliant m elements (i.e. +m, +m– and ms– elements) there is no evidence of over-representation of DR elements. It can generally be concluded, then, that DR elements are no more salient in compliant m elements than non-DR elements.

From Table 6.3, it can also be seen that the subcategory of DR m elements which is most salient, i.e. the category with the highest G^2 figure, is the category of suspensive |m elements. In |m elements, there is almost double the number of DR m elements at 87 compared to the expected frequency of around 45 elements. These can be further divided according to their position in the linear unit, i.e. whether they occur in linear unit-initial position or in a linear unit-medial position. This can be seen in Table 6.4. It can be clearly seen that the over-representation presented in Table 6.3 is due specifically to the over-representation of DR |m elements in an initial position ($G^2 = 37.77$) in the linear unit and not in a medial position ($G^2 = 0.02$).

Table 6.4: DR and non-DR suspensive m elements

	Observed frequencies		**Expected frequencies**		
	DR	Non-DR	DR	Non-DR	G^2
|m (initial)	80	605	38.06	646.94	37.77**
|m– (initial)	26	251	15.39	261.61	6.49

|m (medial)	7	113	6.67	113.33	0.02
|m– (medial)	3	27	1.67	28.33	0.92

It can also be seen in Table 6.3 that there is an over-representation of DR elements in the category of core m elements, albeit of weak significance with a G^2 figure of 10.24 providing some evidence that DR elements are more salient in shorter simple core elements rather than in lengthier cleft core sequences.

We can conclude, then, that DR elements are over-represented to a significant extent in the subcategory of |m elements in initial position. This will form the basis of the discussion of DR m elements in section 6.5.2.

6.3 Subcategories of oi elements

6.3.1 Subcategories of oi elements in general

Research Question 3: *How are oi elements in general – and DR oi elements in particular – distributed in the IMDb corpus discourse?*

Research Sub-question 3a*: What are the relative frequencies of oi elements in terms of their linear compliance, i.e. whether they are core, compliant or suspensive elements?*

As was outlined in Chapter 4, the categorization of oi elements in this study was refined from LUG to distinguish between core oi elements (oi), compliant supplementary oi

elements (ois) and suspensive oi elements (|oi). The frequency of each of these subcategories is presented in Table 6.5.

Table 6.5: OI elements by linear compliance

	Freq.	**%**
oi	127	11.10
ois	298	26.05
|oi	719	62.85
Total	***1144***	***100.00***

The most frequent category is therefore |oi elements at almost 63% of all oi elements. The high relative frequency of suspensive |oi elements is to be expected since, according to Brazil (1995), oi elements are never necessary for the linear unit to reach the target state (see section 4.3.2.7). Instead, in linear terms, they either act as a temporary suspension of the progress towards the target state or are optional supplementary elements at the end of the linear unit when the target state has already been reached. We can see that there are a large number of |oi elements in the IMDb corpus, and a lower number of ois elements. It is also of note that just over 11% of oi elements are core oi elements. As explained in section 4.3.2.7, these tend to be oi elements acting as stand-alone interactional responses to a previous poster's turn (e.g. W*ell said*; *eww*; *You're welcome* etc.).

It is also useful to distinguish between linear unit-initial |oi elements and linear unit-medial |oi elements just as was done for suspensive m elements in Table 6.3. As can be seen in Table 6.6, linear unit-initial |oi elements represent the majority of |oi elements with 676 out of a total of 719 |oi elements, i.e., over 94%. These include such hedges as *I*

think; attitude markers such as *more importantly*; engagement markers such as *I mean* and *no* and so on. Linear-unit-medial loi elements, on the other hand, are quite rare, with only 43 occurrences in the corpus, i.e. slightly under 6% of all loi elements. These include *like*, *the f*, *well*, *mate* and so on.

Table 6.6: Suspensive oi elements by linear unit position

	Freq.	**%**
loi (initial)	676	94.02
loi (medial)	43	5.98
Total	***719***	***100***

In the next section, the question of how DR elements are distributed in the subcategories of oi elements will be explored.

6.3.2 Subcategories of discourse reflexive oi elements

Research Sub-question 3b: *Are DR oi elements distributed in the same way as non-DR oi elements in terms of their linear compliance or are there significant differences in their distribution?*

As was reported in Table 5.4 in Chapter 5, DR elements are over-represented in oi elements in general. Table 6.7 shows the distribution of DR and non-DR elements in oi elements. It can be seen that both core oi elements and loi elements are strongly over-represented with G^2 values of 45.65 and 51.68 respectively. DR elements are also over-represented in compliant ois elements but the over-representation is not significant.

Table 6.7: DR and non-DR oi elements

	Observed frequencies		Expected frequencies		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
oi	30	97	7.06	119.94	45.65**
ois	27	271	16.56	281.44	5.92
loi	91	628	39.94	679.06	51.68**

Just as with the distribution of DR and non-DR suspensive m elements, when these results are further divided, it can be seen in Table 6.8 that it is the linear unit-initial loi elements in which DR elements are over-represented. Hence both initial loi elements and initial lm elements are the most salient DR elements in the IMDb corpus indicating that the high number of initial suspensive DR elements is clearly a salient feature of linear units in IMDb corpus discourse.

Table 6.8: DR and non-DR suspensive oi elements

	Observed frequencies		Expected frequencies		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
loi (initial)	87	589	37.56	638.44	51.22**
loi (medial)	4	39	2.39	40.61	0.97

We can therefore conclude that DR elements are over-represented in loi initial elements and core oi elements. These will form the basis of the discussion in section 6.5.3.

6.4 Subcategories of ot elements

6.4.1 Subcategories of ot elements in general

Research Question 4: *How are ot elements in general – and DR ot elements in particular – distributed in the IMDb corpus discourse?*

Research Sub-question 4a*: What are the relative frequencies of ot elements in terms of their linear compliance, i.e. whether they are compliant or suspensive elements?*

As can be seen in Table 6.9, ot elements have a noticeably different distribution in terms of linear compliance when compared to oi elements. As was explained in Chapter 4, a fundamental difference between ot and oi elements is that ot elements generally serve as a means of moving the linear unit towards its target state. A large number of ot elements are therefore compliant, whilst oi elements are generally considered to be holding up the progress towards reaching the target state and are therefore suspensive. It is unsurprising, then, that the category of compliant +ot– elements is by far the most frequent of the categories of ot elements, as seen in Table 6.9, accounting for over 89% of all ot elements. In contrast to oi elements, the number of suspensive |ot elements is low at 181, i.e. around 7.5% of ot elements. The number of ots elements is also low at around 3.4% of all ot elements.

Table 6.9: ot elements by linear compliance

	Freq.	**%**
+ot–	2142	89.06
ots	82	3.41
|ot	181	7.53
Total	***2405***	***100***

As can be seen in Table 6.10, when suspensive ot elements are divided according to their position in the linear unit, in contrast to |oi elements, |ot elements in a medial position (e.g. *and*, *or* and *however*) are more frequent than in initial position (e.g. *although*, *since* and *because*).

Table 6.10: Suspensive ot elements by linear unit position

	Freq.	%
|ot (initial)	75	41.44
|ot (medial)	106	58.56
Total	***181***	***100***

All of this would seem to indicate that oi elements and ot elements do indeed fulfil very different roles in the IMDb discourse with the most striking result being that almost 90% of ot elements are compliant. As has been seen in Chapter 5 in Table 5.11, the vast majority of these ot elements are extremely high frequency conjunctions such as *and*, *but*, *that* and *or*. In contrast, oi elements in the IMDb corpus discourse are, for the most part, suspensive and are much more diverse. They are not strictly necessary to the linear unit in achieving its target state and are instead interventions on the interactive plane to orient the reader as to how the upcoming linear unit should be approached.

6.4.2 Subcategories of discourse reflexive ot elements

Research Sub-question 4b: *Are DR ot elements distributed in the same way as non-DR ot elements in terms of their linear compliance or are there significant differences in their distribution?*

Table 6.11: DR and non-DR ot elements

	Observed frequencies		**Expected frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
+ot–	54	2088	119.00	2023.00	-46.73**
|ot	2	179	10.06	170.94	-10.02
ots	5	77	4.56	77.44	0.04

The difference between ot and oi elements is also clear when we look at discourse reflexivity. Fundamentally, the number of DR ot elements in all subcategories is very low, particularly in ots and |ot elements. In the case of compliant +ot– elements, the under-representation of DR elements is of a strongly significant nature, meaning that compliant +ot– elements is not associated with discourse reflexivity in the IMDb corpus data. This significant under-representation of DR ot elements is not evident in ots or |ot elements but the figures are so low here that it is difficult to draw firm conclusions. Relating these results back to section 5.4.5, those DR +ot– elements that do exist in the IMDb corpus data tend to be those frame markers (such as cardinal and ordinal numbers) and code glosses (such as *for example*), which generally occur at the beginning of linear units. In other words, DR ot elements tend to provide links between linear units rather than links between elements within the linear unit.

Based on the results in this section, two of the subcategories described above (|m elements in linear unit-initial position and |oi elements in linear unit-initial position) will form the focus of section 6.5.

6.5 Discourse reflexivity in the most salient element subcategories

6.5.1 Introduction

The purpose of this section is to describe the characteristics of the most salient DR elements in the IMDb corpus and how they function in the IMDb corpus discourse. This will be done by proposing a lexico-linear means of describing these elements and the element sequences in which they occur and then by illustrating their role in linear unit relations in the IMDb discourse.

In order to do this, several stages will be followed. Firstly the two subcategories which represent those in which DR elements are at their most salient in the IMDb corpus are chosen, namely:

- |m elements in linear unit-initial position
- |oi elements in linear unit-initial position.

Both of these subcategories will be presented firstly in terms of their lexical composition. In the case of |m elements in linear unit-initial position, this will be done through drawing up a KeyWord comparison between |m elements and all other m elements. In the case of |oi elements in linear unit-initial position, this will be done through a frequency list of elements. Secondly, another KeyWord comparison will be used to compare DR elements with non-DR elements in the subcategory. This, in turn, will yield a KeyWord list with the most salient words in the DR elements of the subcategory. The most salient words from the KeyWord list will then be chosen and a lexico-linear characterization of the elements and element sequences in which they occur will be presented. Finally, the role

of these elements will be explored by illustrating the linear units in which they occur and their role in the IMDb discourse.

6.5.2 Discourse reflexive |m elements in initial position

6.5.2.1 General characteristics

In this section, we will firstly examine the characteristics of |m elements in initial position in general in the IMDb corpus before going on to look at DR |m elements in initial position in particular.

Table 6.12: Frequency WordList of |m (initial) elements

N	Word	Freq.	%
1	I	239	7.67
2	IF	115	3.69
3	YOU	108	3.46
4	THINK	93	2.98
5	THE	73	2.34
6	IS	59	1.89
7	IT	58	1.86
8	HE	54	1.73
9	WHAT	44	1.41
10	A	41	1.31
11	WAS	41	1.31
12	DO	35	1.12
13	TO	34	1.09
14	DON'T	32	1.03
15	HAVE	32	1.03
16	WHEN	32	1.03
17	KNOW	31	0.99
18	HOW	30	0.96
19	SAY	27	0.87
20	THAT	27	0.87

A WordList by order of frequency of |m elements in initial position was drawn up as seen in Table 6.12. Immediately apparent is that the first and second person pronouns *I* and

you play a prominent role in |m elements in initial position in the IMDb corpus discourse, thus confirming the immediate interactive nature of the discourse in the corpus and of these suspensive elements in particular. It is also possible to see that the most common content word is *think* with 93 occurrences.

Table 6.13: KeyWords: |m (initial) vs. rest of m elements

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	IF	115	3.6883	61	0.1852	350.49	3E-17
2	THINK	93	2.9827	27	0.082	334.69	3E-17
3	I	239	7.6652	632	1.9183	270.71	6E-17
4	YOU	108	3.4638	291	0.8833	117.37	1E-15
5	THOUGHT	25	0.8018	18	0.0546	67.353	2E-14
6	WHEN	32	1.0263	39	0.1184	66.23	2E-14
7	DO	35	1.1225	66	0.2003	53.2	1E-13
8	HOW	30	0.9622	47	0.1427	52.617	1E-13
9	SAY	27	0.8659	38	0.1153	50.991	2E-13
10	WHAT	44	1.4112	123	0.3733	45.381	5E-13
11	KNOW	31	0.9942	67	0.2034	41.759	2E-12
12	SAID	19	0.6094	23	0.0698	39.427	5E-12
13	DON'T	32	1.0263	85	0.258	34.928	5E-10
14	WHY	22	0.7056	42	0.1275	33.04	6E-09
15	SURE	11	0.3528	6	0.0182	32.9	7E-09
16	BELIEVE	16	0.5131	25	0.0759	28.068	1E-07
17	GLAD	8	0.2566	3		26.838	2E-07
18	DID	22	0.7056	53	0.1609	26.619	2E-07
19	DOES	17	0.5452	31	0.0941	26.5	3E-07
20	DOUBT	7	0.2245	2		25.114	5E-07
21	TO	34	1.0904	776	2.3554	-25.12	5E-07
22	OF	26	0.8339	665	2.0185	-26.4	3E-07
23	THE	73	2.3412	1355	4.1129	-27.06	2E-07
24	A	41	1.3149	914	2.7743	-28.3	1E-07

Table 6.13 shows the results of the KeyWord list of |m elements in initial position compared to all other m elements in the corpus. We can see that the two most salient words in Table 6.13 are *if* and *think*, the first of which is an open selector introducing the

premise or limitation of a conditional the second of these occurs commonly as part of a V *that* pattern. These can be directly related to the categories of suspensive m elements presented in section 4.3.2.5. The first of these, *if*, can be related to Category d., i.e. an element which begins the linear unit with an open selector; and the second *think* can be related to Category g., i.e. an element which contains a noun or NV sequence but one that acts as a preface to the linear unit. This can be done for all KeyWords in Table 6.13, as seen in Table 6.14.

Table 6.14: KeyWords: |m initial position elements

Cat.[76].	**Description**	**Related KeyWords[77]**
d.	Elements which begin the linear unit with an open selector	*if*; *when*; *how*; *what*; *why*
g.	Elements which contain a noun or NV sequence but one that acts as a preface to the linear unit	*think*, *thought*, *say*, *know*, *said*, *sure*, *believe*, *glad*; *doubt; don't*
h.	Elements which contain a VN sequence rather than a NV sequence.	*do*; *did*; *does*

It can therefore be concluded that the three categories of suspensive m elements in linear unit initial position featured in Table 6.14 are the most salient in the IMDb corpus discourse.

[76] These categories follow the denomination from section 4.3.2.5. There are eight categories in total. Categories a.–c. are categories for compliant elements. The words in Table 6.13 do not easily relate to Categories e. and f., hence categories d., g. and h. only feature in Table 6.16.
[77] The personal pronouns are not included in Table 6.14 as they may be related to all categories.

Having established the characteristics of |m elements in general in initial position, the characteristics of DR |m elements in initial position can be examined. As can be seen in Table 6.15, the two most frequent words in DR |m elements are the first and second person pronouns, *I* and *you*. These therefore play as prominent a role in DR |m elements in initial position as has been established that they do in DR |m elements in initial position in general. The third most frequent word is the discourse verb *say*. It is also of note that *said* and *saying* are the ninth and tenth most frequent word respectively in the table.

Table 6.15: Frequency list of words in DR |m initial position elements

N	Word	Freq.	%
1	I	21	7.5
2	YOU	17	6.07
3	SAY	14	5.00
4	TO	10	3.57
5	IS	7	2.50
6	THE	6	2.14
7	IF	5	1.79
8	IT	5	1.79
9	SAID	5	1.79
10	SAYING	5	1.79
11	ARE	4	1.43
12	AS	4	1.43
13	CAN	4	1.43
14	WAS	4	1.43
15	WHAT	4	1.43
16	WOULD	4	1.43
17	AM	3	1.07
18	HAVE	3	1.07
19	I'M	3	1.07
20	MEAN	3	1.07
21	MEANT	3	1.07
22	MY	3	1.07
23	TELL	3	1.07
24	THIS	3	1.07

Table 6.16 shows the results of a KeyWord search comparing DR |m elements in initial position with non-DR |m elements in initial position. It can be seen that *say* and *saying* are the first and second most salient words in |m elements in initial position.

Table 6.16: KeyWord: DR |m vs. non DR |m elements

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	SAY	14	5	10	0.5139	28.64	8E-08
2	SAYING	5	1.7857	1	0.0514	15.668	8E-05
3	TO	10	3.5714	13	0.668	13.676	0.0002
4	YOU	17	6.0714	42	2.1583	11.313	0.0008

In the next section, the characteristics of DR |m elements in initial position containing *say* and *saying* will be explored. This will be done by proposing lexico-linear characterizations of these elements such that the element subcategory, the KeyWords of the element and how it combines with other elements in a sequence (if appropriate) can be described. There are three parts to the lexico-linear description:

- **Element subcategory**: the subcategory of the element in which the lexical item occurs and its position in the linear unit.
- **Lexical item**: This may be one word or a phrase. Following Cheng et al (2009), a slash is used to indicate that there may be a number of intervening words between the two words. If a slash is not present then it indicates that it is a fixed phrase.
- **Further elements**: Further elements may appear before or after the element in question if there is evidence of a typical element sequence. These additional elements may appear in brackets indicating that they are optional.

For instance, a **|m (initial) *I/SAY* (+ot– *that*)** sequence, indicates that it is a suspensive |m element in initial position in the linear unit, that it includes the words *I* and *say* and that there may be a number of intervening words between *I* and *say*. The bracket indicates that it may or may not be followed by a +ot– element containing *that*.

6.5.2.2 Two meanings of DR |m elements in initial position containing **say** *and* **saying**

It will be argued here that *say* has two broad meanings in |m elements in initial position in the IMDb corpus data. The first of these is related purely to the literal meaning of *say*, i.e. communicating through speaking words. This can be seen for example in a variety of |m elements: *How can you say*; *I'm trying to say* and so on. The *-ing* form *saying* also tends to represent this meaning. Examples of these include: *I'm just saying; you're saying*; *I'm not saying* etc. Such elements are unequivocally discourse reflexive as they clearly refer to a discourse act within the present discourse.

The second broad meaning of |m elements containing *say* in initial position is primarily a signal that the upcoming element sequence is an opinion. This is particularly true if *say* is combined with the first person pronoun *I* as subject. Examples of these include: *I'd say*; *I'd tend to say; I'll have to say; I could say; I also have to say* and so on. These can therefore be characterized as being a **|m (initial) *I/SAY* (+ot– *that*)** sequence. As can be seen from the examples listed above, these elements tend to have some sort of modality in their meaning containing, as they do, a modal or semi-modal verb. These elements can therefore be conceived as providing an indication of the extent of the averral of the writer.

As they do not primarily convey a literal meaning they can be considered to express an 'idiomatic' meaning.

This second type of |m element communicates 'a firm expression of the speaker's personal view' (Sinclair and Mauranen, 2006:78) as well as the presence of the poster and is therefore considered to be an m element. The **|m (initial) *I/SAY* (+ot– *that*)** sequence can be described using the adaptation of Hyland's categories of interactional metadiscourse, as employed in section 5.3.4. Using this categorization, the majority of these elements can be considered to act primarily as a hedge, for example, *I would say*, *I'd say*, *I'd tend to say*, *I'll have to say* and *I also have to say*. The prominence of **|m (initial) *I/SAY* (+ot– *that*)**, as an expression of stance in initial position reflects the prominence of other such non-DR sequences as: **|m (initial) *I/THINK* (+ot– *that*)**; **|m (initial) *I/BELIEVE* (+ot– *that*); |m (initial) *I'M/SURE* (+ot– *that*)** and so on, all of which feature in Table 6.13. The prominence of these types of stance marking sequences containing a |m element would seem to go some of the way to explaining the relative high frequency of linear units containing this type of fronted initial |m element. Such an explicit expression of stance is, after all, what would be expected in message boards whose purpose it is to allow posters the opportunity to express their opinions.

Because of all of the above, it may well be argued that **|m (initial) *I/SAY* (+ot– *that*)** is primarily an expression of stance rather than of a reference to the present discourse and that it therefore may not be discourse reflexive at all. It certainly would not be included, for instance in Ädel's speech event model of metadiscourse (see section 3.4.2). Instead, it

would be categorized as being a writer-oriented stance marker (Ädel 2006:40). It is therefore marginal in terms of fulfilling the criterion of being an explicit expression of a discourse act as stipulated in the definition of discourse reflexivity in this study as seen in section 4.4. This is particularly true of the most fixed of these elements, such as *I'd say* or *I would say*. However, it is also true that the less fixed these elements are, the more one can perceive the literalness of the discourse act. Thus, in elements such as *I'd tend to say*, the physicality of the verb *say* is more prominent as is an increased tone of real time communication. This is seen clearly in the |m– / +m sequence *I'd be / hard-pressed to say* from the IMDb corpus, where the sense of communicating is as prominent as the sense of signalling a stance. The **|m (initial) *I/SAY* (+ot– *that*)** sequence may therefore be characterized as primarily a stance marker but one which also draws attention to the discourse act. It is therefore classified as being discourse reflexive.

6.5.2.3 DR |m elements in initial position containing **say** *and* **saying** *by turn orientation*

In this section, examples will be presented and discussed of |m elements in initial position containing *say* and *saying*. This will be presented via two variables. The first is by meaning. As has been argued above, |m elements in initial position containing *say* and *saying* have two main meanings: the first *say* in a literal sense of telling; and the second as a stance indicator. The second variable to be taken into consideration will be the orientation of the linear units, i.e. whether they have intra-turn or inter-turn orientation. These variables can be summed up in Table 6.17. The descriptions that follow will follow the order (i), (ii), (iii) and (iv) as indicated in the table.

Table 6.17: |m initial position elements containing *say* and *saying*

	Intra-turn	**Inter-turn**
Meaning of *say/saying*: Literal	(i)	(ii)
Meaning of *say/saying*: Stance	(iii)	(iv)

(i) |m elements in initial position containing say/saying *with literal meaning and intra-turn orientation*

In linear units with intra-turn orientation, |m elements containing the literal meaning of *saying* appear in the corpus data in suspensive linear units. They therefore act as a signal that the poster feels that communication has not been successfully completed at this stage. In other words, the poster is communicating, through the use of the element, that the target state has not been reached and that the upcoming linear unit will address this situation. For instance, in Example 6.1 the linear unit in line 2 contains the DR sequence **|m (initial) *I'M/SAYING* (+ot– *that*).** In this case, it acts as a signal that the rest of the linear unit is an overlay with the antonym *inaccurate* contrasting *accurate* in line 1. In non-technical terms, this suspensive element provides a clarification of the previous statement or a means to emphasize the poster's self consistency (Mauranen 2012:177). This is then supplemented by a further concession in line 3. In terms of other similar studies, the linear units in lines 2–3 can be considered to be a code gloss (Hyland 2005) or a clarification (Ädel 2006, 2010). This would therefore seem to be fulfilling Jakobson's metalinguistic function of language referring to itself at moments of clarification as discussed in Chapter 3.

Example 6.1[78]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1.	…just because something is in print doesn't make it accurate.		M
	2.	**I'm not saying /** it is all inaccurate/ either	**|m /** m **/** ots	**|M (overlay)**
	3.	but the only people that know are dead		MS

Similarly, in Example 6.2, the DR element **|m *I'M/SAYING* (+ot– *that*)** in line 4, is used by the poster in apparent anticipation that the reader will not have fully understood the previous linear units, i.e. that the poster feels that communication has not been fully successful up to that point. Hence, the DR |m element in line 4 represents another example of **|m (initial) *I'M/SAYING* (+ot– *that*)** functioning as a signal that the upcoming element sequence is a suspensive clarification of some sort. In the case of Example 6.2, however, the suspension is more complex than that seen in Example 6.1 as it represents the second linear unit in a sequence that extends over two linear units. The suspension is initially signalled by the non-DR |oi element *Don't get me wrong* in line 3. This is then followed by a clarification in the rest of line 3 and continues into line 4. The first of these, in line 3, is a suspensive linear unit. This linear unit seemingly anticipates a misunderstanding on the part of the reader and therefore states what is not the case. However, it also prospects that a statement of what *is* the case will follow. The linear unit in line 3, which contains the DR sequence **|m (initial) *I'M/SAYING* (+ot– *that*)** provides this statement thus clarifying what the poster's actual attitude is. In terms of rhetorical structure (Winter 1977), the DR element here acts in the second part of a matching

[78] This example is from Thread 005 about English actor Cary Grant. Here the poster is criticizing a biography of the actor.

relation expressing the structure: denial and correction. Thus, this poster would seem to be showing a degree of sensibility as to the anticipated problems the reader is likely to have in understanding the message. Discourse reflexivity, therefore, is seen as a means of demonstrating a willingness to anticipate readers' problems. It is an example of discourse reflexivity performing an 'act of consideration' (Mauranen 2010). As such, this can be considered to be a strategy on the part of the poster to succeed in being more effective in persuading the reader.

Example 6.2[79]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1.	I love the original movie and play,		M–
	2.	and I think they would be perfect to remake it starring George Clooney.		+M
	3	Don't get me wrong, I don't think the movie needs remaking at all,		|M–
	4	**I'm just saying,** /that /if they did announce /that /they were working /on this / I would be extactic.	**|m /** +ot– / |m / +ot–/ m– / +m/ m	**+M**

However, the motivation to use suspensive linear units with a DR |m element containing *say*/*saying* with an intra-turn orientation may not always be as transparent as it may seem. Instead of being a means to help readers understand a potentially confusing point, they may be considered to be a means of surreptitiously rephrasing what the poster had previously stated. In Example 6.3, **|m (initial)** ***WHAT I'M TRYING TO SAY/IS*** **(+ot–**

[79] This example is from Thread 049 about American director Joel Coen.

***that*)** in line 6 again can be seen to be acting as a signal that the poster feels that what has been stated before may not have reached the target state satisfactorily. However, in this case the linear unit which contains **|m (initial) *WHAT I'M TRYING TO SAY/IS* (+ot– *that*)** has a second function specifically related to linear structure. It is also used to signal that that the upcoming linear unit does not refer to the immediately preceding linear unit. It is clear that the linear unit in line 6 is not simply a neutral paraphrase of the linear unit in line 5. Indeed, it is not entirely clear what it is a paraphrase of. It may be interpreted as a paraphrase of line 4, of lines 2–4 or even something that the poster had written in a previous post. In the end, though, the reference is vague and is no more specific than letting the reader know that it does not refer, in Sinclair's terms, to the text of the moment but rather to something in the shared information between the poster and the reader somewhere in the records of the autonomous plane. The fact that **|m (initial) *WHAT I'M TRYING TO SAY/IS* (+ot– *that*)** does not refer to the immediately previous linear unit is also made clear through the use of a suspensive DR element sequence: *and / apparantly / **I said it/** in the wrong way*, which itself suspends the original suspensive element. Given this, it would seem most likely that the poster is going to present a rephrasing of something said in a previous post.

Example 6.3[80]

P	L	Linear Unit	DR elements	Linear Unit Type
D	1.	Ok fine, perhaps I was harsh.		MS–
	2.	It of course is not as bad as troll 2 and battlefeild earth.		+M
	3.	It was actualy ok and deserved a 6.5 to a 7.		MS

[80] This example is from Thread 059 about the film *Alien* (see Example 8.1)

	4.	I thought you thought I was the one who posted the MIAMI VICE WORST MOVIE EVER post		M
	5.	so that is why I was a little irritated.		MS
	6.	**What I am trying to say,** / and / apparantly / **I said it/** in the wrong way / is / that / I think / this movie is good, /but / overrated.	**|m– /** |ot / |oi / **|m** / ms / +m– / +ot– / |m/ m / +ot– / ms	**|M**

As will be seen in Example 8.1a in Chapter 8, the poster had in fact initiated the thread with a very critical post regarding the film in question, which s/he refers to as 'the most overated [sic] sci fi of all time' and that it 'belongs with battlefeild earth and troll 2' (two films which are widely seen as being of very low quality). The poster has already rescinded the second of these statements in line 3 in Example 6.3. Thus when the reader reaches ***What I am trying to say, / and / apparantly / I said it/*** *in the wrong way / is / that* in line 6, s/he has contradictory expectations. On the one hand, s/he expects the upcoming linear unit to be a felicitous rephrasing of the original opinion that it is the most overrated film of all time, on the other hand, s/he is aware that the poster has rescinded his/her earlier negative statements about the film. What is provided in the rest of the linear unit in line 6 is something that looks like a felicitous reiteration of the previous statement, through the repetition of the word *overated* [sic]. However, the poster has surreptitiously included a positive evaluation of the film: *this move is good*, which was not present in the previous turn. It is therefore an example of the use of a DR |m element being used to supposedly signal clarification but in fact it is a tool to backtrack on a contentious previous statement. As Mauranen (2012:178) comments, such elements are 'a useful means of promoting the [poster's] position and making it *appear* consistent' (my italics).

The poster seemingly feels that s/he can achieve this precisely because s/he is relying on the fact that what exactly was previously stated remains in the shared records of the autonomous plane and that other participants are unlikely to remember exactly what was written and/or to go back to check.

(ii) |m elements in initial position containing **say/saying** *with literal meaning and inter-turn orientation*

There is some evidence in the IMDb corpus discourse that linear units with inter-turn orientation that begin with a |m containing *say*/*saying* with the literal meaning of *say*, are generally suspensive linear units just as they are when they have intra-turn orientation. This is particularly true if *say*/*saying* is combined with the subject *you*. For instance, in Example 6.4, **|m (initial)** ***YOU ARE /SAYING*** **(+ot–** ***that*****)** would seem to be signalling a request for a clarification.

Example 6.4[81]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1.	What has happened since then is the result of Tom mistakenly believing that he is so well-loved that he could speak openly about his freakish cult without suffering any backlash.		MS
	2.	He was wrong.		MS
B	3.	So / basically, /**you are saying** /that /its wrong to speak out /about your beliefs?	+ot– / |oi/ |**m** / ot/ m/ ms	**|M** (overlay)

[81] This example is from Thread 022 about American actor tom Cruise.

As we have seen with |m elements in initial position containing *say*/*saying* with literal meaning and intra-turn orientation, at first sight this would seem to be a request for a code gloss (Hyland 2005) or a clarification (Ädel 2006) through the use of an elicitation. However, on closer reading, it is clear that Poster B does not expect Poster A to respond to the linear unit in line 3. It would therefore seem to be inaccurate to say that it prospects a response. Instead, it is better conceived as a means of criticizing the previous poster's statements in Lines 1 and 2. In the terminology of this study, this is done through the use of an overlay signalled by the combination of |oi *Basically* and **|m (initial) *YOU ARE /SAYING* (+ot– *that*).** The overlay repeats the word *wrong* (in line 2) and rephrases the previous linear unit in line 1. However, it is clear that this is not a genuine attempt to rephrase the previous linear unit to create a linear unit which is synonymous with the previous post. Instead, it is an aggressive attempt to willfully misinterpret what was written in order to highlight its supposed absurdity. In non-technical terms, Poster B is putting words into Poster A's mouth.

A similar example can be seen in line 2 in Example 6.5, where the poster also appears to be making a request for clarification in an interrogative form. However, in reality, it is more accurate to say that it is a signal of indignation towards the previous poster's assertion. The sequence **|m (initial) *HOW CAN YOU SAY* (+ot– *that*)** signals either a verbal echo or overlay, depending if the poster chooses to paraphrase or quote exactly. In this case it is a paraphrase and so can be categorized as an overlay. The linear unit in line 2 in turn prospects a development of the stance to explain why it cannot be said that the film in question is the worst Oscar movie ever. This is provided in line 3.

Example 6.5[82]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1.	The Worst 'Best' oscar movie ever????		M
B	2.	**How can you say** / this is /the worst oscar movie ever	**|m** / m– /+m	**|M–** (overlay)
	3.	when you have movies like Titanic, Shakespeare in Love, and Lord of the Rings that won oscars for best picture?		+M

This type of aggressive rephrasing signalled by a DR element through an overlay is one example of how DR elements are employed in the IMDb corpus as tools used in points of disagreement and conflict. This will be explored further in Chapter 8.

(iii) |m elements in initial position containing say/saying *with a stance meaning and intra-turn orientation*

When we examine |m elements in initial position containing *say*/*saying* with stance meaning and intra-turn orientation, there is similar evidence that DR elements play a part in the tone of antagonism evident in the posts. In previous studies of hedges (e.g. Hyland 1998a, 2000), such sequences as **|m (initial) *I/SAY* (+ot– *that*)** are presented in terms of being means of signalling tentativeness and politeness by limiting the commitment of the writer to the upcoming proposition. In the IMDb corpus discourse, however, not all of the DR **|m (initial) *I/SAY* (+ot– *that*)** sequences simply fulfil such a role. As stated

[82] This example is from Thread 067 about the film *Braveheart*.

previously, much of the tone of the IMDb corpus is ironic and this is reflected in some of the DR **|m (initial) *I/SAY* (+ot– *that*)** sequences in the IMDb corpus data. For instance, in Example 6.6, there is an ironic juxtaposition between the tentative and the assertive in close proximity.

Example 6.6[83]

P	L	Linear Unit	DR elements	Linear Unit Type
	1.	As for me, / **I'll tell you plainly /** what I think :	|oi /**m–** / +m	**M–**
	2.	**2) / I'd tend to say**/ that/ the "poor lambs"/ in question / are goddamn loonies / to despise the interest / such a great actor / takes in them	**+ot–** / **|m** / +ot– / m– / |m/ +m / ms– / +m– /+m	**MS**

The assertiveness is provided in the preamble leading to the expression of stance in line 1. This prospects the rest of the turn (of which only a part is reproduced here) and includes an initial |oi element, *As for me,* which indicates to the reader that the upcoming series of elements should be taken as a person opinion. Line 1 also includes a DR element which prospects the upcoming elements with the DR verb *tell*. This is modified by the adverb *plainly*, pre-warning the reader that forthright and possibly unpalatable opinions will follow. Further assertiveness is seen in the actual opinion and the use of the ironic quotation marks *"the poor lambs"* in line 2 and the ironic contrast between calling these people *poor lambs* and *goddamn loonies*. The supposed tentativeness is represented by the **|m (initial) *I/SAY* (+ot– *that*)** sequence in line 2: *I'd tend to say*. From this, the reader

[83] This example is from Thread 013 about American actor Kevin Spacey (see Example 8.2).

would expect the upcoming elements to be introducing a tentative opinion. As we have seen, what follows in the rest of the linear unit is in fact a very forthright opinion, bordering on being insulting. The fact that what is introduced is so opinionated creates an ironic sense that if the poster had chosen to the tell the truth more plainly rather than supposedly hedging it as s/he signals with *I'd tend to say*, then, the opinion expressed would have been even more extreme. The use of **|m (initial) *I/SAY* (+ot– *that*)** is therefore used as a means of ironic juxtaposition here to heighten the sense of irony in the criticism of the previous poster's turn.

There is also evidence that |m elements in initial position containing *say*/*saying* with stance meaning and intra-turn orientation play a specific role in longer turns. In this context these elements actually signal two things. They act as a means of signalling that a summarizing reiteration of the position stated in the turn is coming up and they also indicate that the upcoming linear unit is relevant to a point further back in the previous discourse than the immediately preceding linear unit. This can be seen in Example 6.7. In the example, the DR **|m (initial) *I/SAY* (+ot– *that*)** features in line 12 after a reasonably extended explanation in lines 4–11 responding to a question regarding Henry Fonda's attitude towards his daughter, Jane's, anti-war protests presented in lines 1–3.

Example 6.7[84]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1.	Jane's anti War and America behavior		M
	2.	How do you think he took it		M–

[84] This example is from Thread 008 about American actor Henry Fonda.

		when heard all of what Jane was doing against the Vietnam war and all the dislike she got from others, being called a traitor.		
	3.	As a father how do you think he took it?		|M–
B	4.	Henry was a very patriotic man		M–
	5.	and he enlisted to fight in WW2		+M
	6.	originally the military wanted him to be an entertainer,		MS–
	9.	but Henry desired to do the most he could for his country		+M
	11.	and he got a "real" job fighting for the country.		MS
	12.	**I would say** / that / Henry definitly disagreed / with Jane /and / her radical Anti-American beliefs.	**|m /** +ot– / m– / +m / +ot– / ms	**M**

In the example, the Linear Unit in line 12, which contains the DR **|m (initial)** ***I/SAY*** **(+ot–** ***that*****)** sequence, is used as a sort of reiterative stand-alone coda at the end of the turn. The use of this sort of stand-alone coda, i.e. a core M Linear Unit, is a frequent means by which an extended turn is completed in the IMDb corpus discourse. These codas act as both a sort of reiteration or restatement of the opinions expressed but at the same time are a sort of statement apart from the rest of the turn, evident in the fact that there is no overt retrospective supplementary mechanism present. In this case, **|m (initial)** ***I/SAY*** **(+ot–** ***that*****)** may be seen as signalling that the rest of the linear unit is this type of coda. However, it also functions, as we saw in Example 6.7, as a signal that the relevance of the upcoming linear unit is not to be found in the linear unit immediately previous to the text of the moment, i.e. in the current and previous linear unit, but rather that it is related to something which is now in the shared records of the participants in the

autonomous plane. In non-technical terms, it is in line 12 that the poster finally gets round to answering the questions posed in Lines 2 and 3. The **|m (initial) *I/SAY* (+ot– *that*)** sequence therefore signals to the reader that the upcoming linear units is of optimal relevance to some unfinished business in the shared records of the discourse so far.

Similarly, linear units which contain the **|m (initial) *TO SAY* (+ot– *that*)** sequence in initial position occur three times in the corpus, each time they occur at the end of a reasonably lengthy turn and two of them start with the +ot– elements *so* and *but*. They too act as a signal of an upcoming summary and as a signal that the immediately preceding linear unit is not the most relevant to the linear unit. Instead, the **|m (initial) *TO SAY* (+ot– *that*)** sequence acts a means of linking the upcoming linear unit to something stated previously by another poster. This can be seen in Example 6.8, where **|m (initial) *TO SAY* (+ot– *that*)** signals that the linear unit in line 4 functions beyond simply being a reiteration of the linear unit in line 3. It also signals to a previously made point beyond the text of the moment in line 3, i.e. in line 1. The linear unit in line 1 is itself an inter-turn overlay of a previous poster's comments. Therefore the linear unit in line 4 here ultimately refers to a previous event in the discourse which occurred in a previous post.

Example 6.8[85]

P	L	Linear Unit	DR elements	Linear Unit Type
D	1.	of course they ruin their legacy,		|M–
	2.	when they ham up in rubbish films they make themselves		+M

[85] This example is from Thread 017 about Robert De Niro (see Appendix 4.3).

		look like fools,		
	3.	when Olivier did Inchon-Jazz Singer etc he was absolutely making a fool of himself and / looked ridiculous.		MS
	4.	so **/ to say** / you dont ruin your legacy / is rubbish	+ot– **/ |m /** m– /+m	**MS**
	5.	because you do.		MS

There is also evidence that **|m (initial) *TO SAY* (+ot– *that*)** does more than signalling that the upcoming linear unit has its relevance beyond the immediate text of the moment. From the limited evidence from the IMDb corpus, it would seem that it forms the following sequence:

|m (initial) *TO SAY* (+ot– *that*) + prev. poster's stance **+** negative eval. of that stance

Table 6.18: |m (initial) *TO SAY* (+ot– *that*) sequence

Ex.	**(+ot–)**	**|m (initial) *TO SAY***	**(+ot– *that*)**	**Prev. poster's stance**	**negative eval. of stance**
6.8	*so*	*to say*	**~**	*you don't ruin legacy*	*rubbish.*
6.9	~	*To say*	*that*	*Coppola is overated* [sic]	*to live your entire life in a black box*

For instance, in Example 6.8, the linear unit in line 4 can be described in these terms as seen in Table 6.18. Similarly in Example 6.9, the linear unit containing the **|m *TO SAY* (+ot– *that*)** would seem to be acting in three different ways: firstly as a reiteration; secondly, as a signal that the upcoming element sequence is relevant is to a statement beyond the linear unit immediately previous; and thirdly, that the upcoming evaluation is likely to be negative.

Example 6.9[86]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	That makes him a great director.		MS
	2	to deny this is futile.		MS
	3	**To say** / that / Coppola is overated / is to live your entire life / in a black box.	**|m** / +ot– / m– / +m– /+m	**M**

Thus, again, DR elements would seem to play a central role in expressing antagonism and conflict. This will be further explored in Chapter 8.

(iv) \|m elements in initial position containing say/saying *with stance signal meaning and inter-turn orientation*

Linear units beginning with a |m element with *say* where it is a stance marker, tend to occur in compliant linear units when they have inter-turn orientation. They do not tend to possess the same complexities of use as has been described above in the other three categories. This is most clearly seen when it occurs in a +M Linear Unit with inter-turn orientation. For instance, in Example 6.10 the linear unit in line 2, which contains a **\|m (initial)** ***I/SAY*** **(+ot–** ***that*****)** sequence simply responds to the initial elicitation posted by the original poster.

Example 6.10[87]

P	L	Linear Unit	DR elements	Linear Unit Type

[86] This example is from Thread 052 about American director Francis Ford Coppola.
[87] This example is from Thread 018 about American actor Robert De Niro.

A	1.	Roberts Freakiest Role????		M–
B	2.	**Id say** /either /Cape Fear /or / This boys life.	**|m** / +ot– / m / +ot– / ms	**+M**

Therefore, it has been seen here that |m elements in initial position containing *say*/*saying* perform a variety of roles in the linear structure of the discourse. These range from occurring in suspensive linear units used to signal clarification; to compliant linear units at the end of turns indicating a summary and their relevance previous in the discourse; to simply signalling a response in a compliant M– /+M elicitation response sequence.

6.5.3 Discourse reflexive |oi elements in initial position

6.5.3.1 General characteristics

In this section, the characteristics of |oi elements in initial position in general in the IMDb corpus will first be examined before going on to look at DR |oi elements in initial position in particular.

As was explained in Chapter 4, a |oi element which occurs in an initial position in a linear unit flouts the fundamental expectations that a linear unit will begin with an element which will immediately aid in the movement towards the target state. On the contrary, |oi elements in initial position orientate the reader by providing a 'preliminary intimation' (Brazil 1995:184) of how the reader should approach the linear unit without moving the linear unit any nearer to reaching its target state.

In order to investigate the characteristics of DR |oi elements in initial position, a Frequency Word List of element types was drawn up. Secondly, the characteristics of DR |oi elements in particular were established by comparing DR |oi elements with non-DR |oi elements in a KeyWord search.

The Word List by order of frequency of |oi elements in initial position can be seen in Table 6.19. These can be related to the non-structural discourse functions introduced in section 5.3.4. The most frequent |oi element is the hedge *I think.* Other hedges such as *maybe*, *I guess* and *perhaps* also feature in the list. The second most frequent |oi element in the IMDb corpus data is the engagement marker *well.* Other *engagement markers* are also prominent such as the use of the poster's name, *yeah*, *yes* and *now*.

Table 6.19: Frequency WordList of |oi (initial) elements

|oi element	**Freq.**	**%**
I think	51	7.54
well	38	5.62
maybe	25	3.70
Yea / yeah	25	3.70
<name>	23	3.40
yes	19	2.81
now	17	2.51
no	13	1.92
of course	13	1.92
ha / haha / hehe	11	1.63
I guess / guess	11	1.63
Wow	11	1.63
perhaps	10	1.48

at least	9	1.33
i mean	9	1.33
o /oh	9	1.33
Hey	8	1.18
Hmm / hum	8	1.18
Ok / okay / oh okay	8	1.18
sorry	8	1.18
anyway /s	7	1.04
apparently	7	1.04
lol	7	1.04
in fact /infact	6	0.89
wait / no wait/ ok wait	6	0.89

6.5.3.2 Discourse reflexive |oi elements

From the KeyWord list in Table 6.20 it can be seen that the most salient words in DR |oi elements in initial position compared to non-DR |oi elements in initial position are discourse act verbs, namely, *guess*, *mean*, *say*, *admit* and *wait*. Also of interest in Table 6.20, is the acronym *LOL.*

Table 6.20 DR |oi (initial) vs. Non-DR |oi elements in initial position

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	GUESS	11	4.7826	0		37.395	2E-11
2	MEAN	10	4.3478	0		33.959	3E-09
3	SAY	9	3.913	0		30.53	3E-08
4	LOL	8	3.4783	0		27.109	2E-07
5	WAIT	6	2.6087	0		20.288	7E-06
6	ADMIT	4	1.7391	0		13.496	0.0002
7	HERE	4	1.7391	0		13.496	0.0002
8	TO	13	5.6522	14	1.3944	12.411	0.0004

6.5.3.3 DR |*oi elements in initial position containing* **guess, mean** *and* **say**

What we can see both from the WordList in Table 6.19 and the KeyWord list in Table 6.20 is that the most salient types of DR |oi elements are those that contain the words *guess, mean* and *say*.

As we have seen in Chapter 5, the **|oi (initial)** ***I GUESS*** **(+ot–** ***that*****)** sequence is generally considered to be a hedge. However, when we look at the role of the **|oi (initial)** ***I GUESS*** **(+ot–** ***that*****)** sequence in the linear structure of the discourse we can see that it functions differently depending on its turn orientation, i.e. if it is intra-turn or inter-turn oriented.

Example 6.11[88]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	The sad part is Borat had promise but caved in to that JACKASS mentality.		MS
	2	Oh well **/ I guess** / you could call this/ BROKEBACK BORAT.	|oi / **|oi** / m– / +m	**MS** (Encap.)

Firstly, then, the **|oi (initial)** ***I GUESS*** **(+ot–** ***that*****) sequence** in linear units with intra-turn orientation would seem to occur in compliant linear units just as we have see above with **|m (initial)** ***I/SAY*****(+ot–** ***that*****)** sequences with stance meaning and intra-turn orientation. Hence, in linear units with intra-turn orientation, the **|oi (initial)** ***I GUESS*** **(+ot–** ***that*****)** sequence acts as a sort of signal of an upcoming summarizing conclusion or reiteration. In Example 6.11, the **|oi (initial)** ***I GUESS*** **(+ot–** ***that*****)** sequence is used as a means of

[88] This example is from Thread 058 about the film *Borat*.

introducing a humorous ending to a turn. The humour is also enhanced by the |oi element *Oh well* which places distance between the linear units in lines 1 and 2 but at the same time creates a tone of spontaneity as if the poster is thinking of a conclusion as s/he writes. Line 2 is therefore a summary of the rest of the previous turn but a humorous one and this is signalled in part by the **|oi (initial) *I GUESS* (+ot– *that*)** sequence. This would seem to be similar to the role of |m elements containing *say/saying* in Examples 6.7–6.9. Conversely, **|oi (initial) *I GUESS* (+ot– *that*)** in intra-turn orientation can also be seen in more antagonistic contexts. For example, a similar ironic juxtaposition between tentativeness and antagonism, as was seen in Example 6.6, can be seen in the use of **|oi (initial) *I GUESS* (+ot– *that*)** in Example 6.12. In this example, **|oi (initial) *I GUESS* (+ot– *that*)** is used as a means of introducing an insult as if it is some sort of logical consequence of the previous linear unit creating a sense of incongruity between the supposed politeness and rationality of the **|oi (initial) *I GUESS* (+ot– *that*)** and the insult that follows.

Example 6.12[89]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	I was wondering if he's had any other roles as assassins.		M–
B	2	Willia always plays a cop, hitman or someone in the military.		+M
	3	**I guess** / it's because / he's really / a thoroughly mediocre actor, / like / all Republicans.	**|oi** / |ot / m– / +m / ot / ms	MS (Encap.)

[89] This example is from Thread 003 about American actor Bruce Willis.

In contrast to **|oi (initial) *I GUESS* (+ot– *that*)** sequence with intra-turn orientation, it is of note that all linear units with an inter-turn orientation which include **|oi (initial) *I GUESS* (+ot– *that*)** in the IMDb corpus discourse are found to be suspensive. In Example 6.13, for instance, in non-structural terms, **|oi (initial) *I GUESS* (+ot– *that*)** seemingly acts as a hedge for the rest of the linear unit in line 3. However, in the terms of the linear structure analysis of this study, **|oi (initial) *I GUESS* (+ot– *that*)** acts as a signal that the upcoming linear unit is likely to be a suspensive linear unit. In this case, the suspensive linear unit is in the form of an overlay, paraphrasing the linear unit in Line 2. However, significantly, Poster B has changed the original wording so that *all homos* substitutes for *you*. In doing so, Poster B is including Poster A as one of all the *homos*. It therefore represents a case of an aggressive rephrasing of the original language. In non-technical terms, it is an insult, fronted by the hedge *I guess.*

Example 6.13[90]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	It was hilarious.	m	M
	2	I was laughing/ like/ a hyena	m/ ot/ ms	MS (Overlay)
B	3	**I guess** / all homos were laughing / like / hyenas	**|oi** / m/ ot / ms	|M (Overlay)

Although hedges are often used to introduce unpalatable information, here the potential politeness of the hedge is nullified by the insult thus creating irony from this incongruity in a similar way to the use of *I'd tend to say* in Example 6.6.

[90] This example is from Thread 058 about the film *Borat*.

The second most salient word in |oi elements is *mean* with the most frequent element including mean being **|oi (initial)** ***I MEAN***. The uses of *I mean* have been well documented (e.g. Schiffrin 1987; Mauranen 2012). In terms of its role in the linear structure, it important to note that *I mean* exclusively occurs within one participant's turn, i.e. there are only examples of |oi *I mean* occurring in linear units with intra-turn orientation. In Example 6.l4 the **|oi (initial)** ***I MEAN*** can be seen to be a means of signalling an exemplification of the point being made in Line 1. However, as much as signalling an exemplification, it would seem that the use of **|oi (initial)** ***I MEAN*** here is as much to do with providing a sense of interaction and spontaneity traditionally associated with spoken discourse. This can further be seen by the fact that *I mean* is only one of five other oi elements in the same linear unit (*really*, *like*, *i think* and *soo*).

Example 6.14[91]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	The list on this site / should really be taken / without rank / really		M
	2	**i mean** / it has / like / Finding Nemo / i think / on it / soo…	**|oi** / m– / |oi / +m/ |oi / +m / ois	**MS**

This apparent appealing to the collegiality of the readers of the post is made explicit in Example 6.15, where *I mean* is followed by the oi element *come one* [sic], which constitutes both a direct appeal to the readers to acknowledge the obviousness of the original statement and of the veracity of the elaboration.

[91] This example is from Thread 070 about the film *Fight Club*.

Example 6.15[92]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	Who the hell / hired him / for Godfahter,		M
	2	**I mean** /come one / the original cut was 2 hrs.	**loi** / loi /m	**MS**

Unlike *guess* and *mean*, loi elements that contain *say* represent a wide range of element types. These vary in terms of their non-structural discourse function, ranging from a hedge: *it's possible to say*; a booster: *I say,*; an attitude marker: *It kills me to say it*,; to an engagement marker: *Let's say*. They also vary in terms of the type of linear sequence in which they occur. They can be directly followed by an m element (i.e. with no intermediate +ot– element), such as *as you say*; they can be followed by an optional +ot– *that* element, such as, *it's safe to say*; or they can require the +ot– *but* element: it *kills me to say it*. This is summarized in Table 6.21.

Table 6.21: Structure of Linear units containing loi (initial) elements with *say*

	loi	**loi (+ot– *that*)**	**loi +ot– *but***	**Discourse Function**
it's possible to say		√		Hedge
I say,	√			Booster
it's safe to say		√		Booster
As I often say,	√			Attitude Marker
It kills me to say it,			√	Attitude Marker
(What do I say:	√			Engagement Marker
as you say,	√			Engagement Marker
Let's say		√		Engagement Marker

[92] This example is from Thread 052 about American director Francis Ford Coppola.

Unsurprisingly, given this diversity, their roles in the inter-linear unit structure of the discourse also vary greatly. This can again be divided into linear units with an intra-turn and inter-turn orientation. In linear units with an intra-turn orientation, there is some evidence that linear units containing a |oi element with *say* occur in a conclusion or summary as we have seen is the case for linear units containing **|m (initial) *I/SAY* (+ot– *that*)** sequence with a stance meaning and **|oi (initial) *I GUESS* (+ot– *that*),** as described above. For example, in Line 4 in Example 6.16, the linear unit acts as a conclusion after a quite lengthy response, of which only some is reproduced here, to the original question in Line 1. On the one level it can be argued that the poster is hedging a bold statement with two hedges in sequence: *maybe* ***/ it's possible to say***. However, in terms of the linear sequence, as we have seen above, the DR element signals that the reader should refer to the shared records gleaned from taking part in the rest of the discourse rather than the immediately previous linear unit. In this case, the most relevant records that the reader is expected to make a connection with are the subject of the thread in line 1.

Example 6.16[93]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	In your Opinoin what is Tom Hanks worst peformance		M–
B	2	I think no matter how much you look back at Tom reviews, reviewers might have bad-mouthed the movie,		M–
	3	but they never really bad-mouthed Tom's acting ability		+M

[93] This example is from Thread 037 about American actor Tom Hanks.

	4	so / maybe / **it's possible to say** / that/ Tom has never acted badly	+ot– / loi / **loi** / ot /m	**MS**

Similarly, in Example 6.17 the |oi element containing say: **|oi (initial)** ***I SAY*** acts, in Hyland's terms as a booster, but again functions in the linear structure of the discourse at the end of a series of linear units in the turn. The tone of the turn is of indignation. It begins in Line 1 with a suspensive verbal echo of a statement made by a previous poster. The **|oi (initial)** ***I SAY*** is contained in a stand-alone coda ***I say*** *quit trying to stop time and just get on with it,* i.e. a linear unit with no structural connection to the previous one. It may be argued that the **|oi** ***I SAY*** acts to signal some sort of summary of the sentiment expressed in Lines 1–7. However, it can also be said that the **|oi (initial)** ***I SAY*** serves as a means of placing a distance between the linear units in line 10 with that in Line 9. It can therefore be thought of as a reiteration of the stand-alone nature of the coda.

Example 6.17[94]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	Ageing badly?		|OI
	2	I didn't know there was such a thing!		MS
	3	Brad looks incredibly handsome.		M–
	4	He has that look in which he does look his age and looks as if he has gained some wisdom along the years but still has much to learn.		+M
	5	**Take George Clooney / for example,**	**m /ots**	**MS–**
	6	he's only what two years or so		+M

[94] This example is from Thread 027 about American actor Brad Pitt.

		older than Pitt?		
	7	Still looks amazing for his age!		MS
	8	As for Dicaprio, he's showing not his age per se, but more his maturity.		M–
	9	Couldn't help but be proud after I saw him in the Aviator.		+M
	10	**I say**, /quit trying to stop time/ and / just get on with it.	**|oi** / m / +ot– / ms	**M**
		Age is age, it happens to all.		M

Those linear units containing |oi elements with *say* in initial position with an inter-turn orientation perform slightly different structural roles in the IMDb corpus discourse than those described above. It is found that |oi elements containing *say* occur both in compliant and suspensive linear units with inter-turn orientation. A compliant linear unit containing the **|oi (initial)** ***IT KILLS ME TO SAY IT*** **+ot–** ***but*** sequence can be seen in Example 6.18, for example, in what appears initially to be a refusal to respond to the elicitation in the subject line. However, it is actually part of a rhetorical device which feigns refusal to answer as a means of praising the question as being difficult but worthwhile. The suspension is only momentary as the response is provided within the linear unit in Line 2.

Example 6.18[95]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	Name Kubrick's weakest movie.	m	M–
B	2	**It kills me to say it**, / but / I'll go with Eyes Wide Shut.	**|oi** / +ot– / m	+M

[95] This example is from Thread 055 about American director Stanley Kubrick.

Linear units featuring |oi elements containing *say* are also found in suspensive inter-turn Linear Units. Example 6.19 includes the **|oi (initial) *LET'S SAY* (+ot– *that*).** The sequence functions as a signal that the upcoming liner unit is suspensive. This would seem to be motivated by the perceived inappropriateness of the previous contributions and a re-definition of the parameters of the thread. Discourse reflexivity is therefore being used here as a means of stepping out of the discussion of the subject matter in order to manage the thread. This would seem to corroborate findings regarding the use of discourse reflexivity in dialogic contexts (Tanskanen 2007; Mauranen 2012) and its role in the management of discourse.

Example 6.19[96]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	Name Kubrick's weakest movie.	m	M–
B	2	Fear and Desire!	m	+M
A	3	**Let's say /** we discount everything / before 'Paths of Glory'	**|oi /** m /ms	**|M** (Encap.)

6.5.3.4 DR |oi elements in initial position containing LOL

As described in section 5.4.2, *LOL* is the most frequent of several CMC-specific acronyms which contain a reference to the discourse act *laugh*. These include: *LOL* (laughing out loud) and its variants; *ROTFL* (roll on the floor laughing); *LMAO* (laughing my ass off) and the compound acronym *ROTFLMAO* (roll on the floor laughing my ass off).

[96] Ibid.

LOL is the most frequent of these in the IMDb corpus discourse and perhaps because of this it is the most complex. This can be evidenced by the fact that, as we have seen in Chapter 5, it is subject to orthographical variation within the IMDb corpus (*LOL*, *lol*, *Lol, LOOOOOOOOOOOOOOOOOOOL*). Its role in the linear structure of the IMDb discourse is similarly varied. As such, it features as all three types of oi element, i.e. as a compliant ois, as a loi and as a core oi element.

Example 6.20[97]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	I am a heterosexual woman		M
	2	however / with Angelina / I will make an exception, / **lol** / ;)	+ot– lm / m / **ois** /ois	**MS**

Firstly, as an ois element, it occurs at the end of linear units as seen in Example 6.20. This would seem to represent the structural role which most obviously links LOL to it being an attitudinal marker and, what has been called in other studies, a 'compensatory mechanism', i.e. as a means of compensating not having the same paralinguistic clues in online communication as face-to-face communication has (see Arendholz 2013 for a review of the literature espousing this view). Significantly, *lol* here occurs in a series of two ois elements preceding the emoticon *;)*. Almost all emoticons in the IMDb corpus discourse act as ois elements at the end of linear units and *lol* in this case can be seen as providing a sort of attitudinal reiteration of the turn in the same way that an emoticon does.

[97] This example is from Thread 025 about American actress Angelina Jolie.

However, in intra-turn orientation, there is evidence that **|oi (initial)** ***LOL*** acts not only as a means to expression emotion but also to link the preceding and the following linear units. In Example 6.21, **|oi (initial)** ***LOL*** links the linear units in 1 and 2. This would seem to function in a similar fashion to **|oi (initial)** ***I MEAN***, as seen above, in that it signals an elaboration of the previous point albeit with the additional attitudinal aspect of it being a humorous addition.

Example 6.21[98]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	I agree big time / with <name>.		MS
	2	**LOL /** The most recent movies / that DeNiro have been making / had been a joke!	**|oi /** m– / |m / +m	**MS**

Additionally, there is some evidence in the IMDb corpus data that the use of **|oi (initial)** ***LOL*** in M linear units with an inter-turn orientation is also motivated by phatic considerations rather than simply as a means of expressing humour. As indicated in previous studies (Tagliamonte and Denis 2008; Baron 2009), despite the fact that this is a relatively new addition to the language, the literal discourse reflexive meaning of *LOL* has already been superseded by a pragmaticalized function, i.e. has undergone 'pragmatic re-orientation' (Butler 2008). Baron refers to *LOL* in instant messaging, not as an expression of humour but rather as 'a phatic marker', i.e. communicating a shared emotion or as a means of maintaining social contact (Richards et al 1992). This type of

[98] This example is from Thread 018 about American actor Robert De Niro.

phatic use of *LOL* can be seen in the IMDb corpus data. For instance, in Example 6.22, Poster B repeats *LOL* in line 2, as used by Poster B1 in Line 1.

Example 6.22[99]

P	L	Linear Unit	Elements	Linear Unit Type
A	1	**LOL**–/ I think **/ the OP needs glasses.**	**|oi** / |oi / **m**	**M**
B	2	**LOL…/** yeah,/ I agree	**|oi** / |oi / m	MS

However, Poster B goes beyond the repetition of simply **|oi (initial)** ***LOL*** in his/her response. Instead, s/he follows the same element structure for the whole linear unit: |oi / |oi / m.

|oi	**|oi**	**m**
LOL–	I think	**the OP needs glasses.**
LOL…	yeah,	I agree

Such linear structural parallelism would seem to express Poster B's desire to express congeniality with Poster A at least as much as expressing humour.

Similarly, core oi elements containing *LOL* in OI linear units can be seen as expressing a sense of congeniality through repetition. In Example 6.23, the original poster asks the question if anyone thinks he looks like American actress Uma Thurman and posts a link to a photo. Poster B responds to the initial elicitation in line 1 and Poster C follows up with a further response. As can be seen, Poster C here chooses to express interactional

[99] This example is from Thread 039 about American actress Uma Thurman.

support for Poster B's response by choosing the core oi element *LOL*, which Poster B had used in the response thus maintaining the same tone and enhancing a sense of it being a shared common sentiment in the group.

Example 6.23[100]

P	L	Linear Unit	DR elements	Linear Unit Type
A	1	A guy that looks like UMA?	m	M–
B	2	HAHAHAHAHAHHAHA / **LOOOOOOOOOOOOOOOOOOOOL** / ha ha ha ha / jut like uma	ǀoi / ǀ**oi** / ǀoi / m	-M
C	3	**LOL!!!!!!!**	**oi**	OIS

Significantly, however, such oi elements are also used in the IMDb corpus discourse to signal antagonism and conflict as well as conviviality. In Example 6.24, for instance, **ǀoi (initial)** ***LOL*** is used in a suspensive linear unit with inter-turn orientation. In this instance, **ǀoi (initial)** ***LOL*** is employed as a precursor for an insult. In non-technical terms, **ǀoi (initial)** ***LOL*** in this case becomes an ironic laugh *at* the other posters rather than a jovial laugh *with* the other posters as was the cases in Examples 6.22 and 6.23.

Example 6.24[101]

P	L	Linear Unit	Elements	Linear Unit Type
A	1	It just saddens me to find that people like him are really NOT that cool or compassionate as they seem.		MS
B	2	**Lol, / you americans are / so weak-minded**	ǀ**oi** / m– /+m	ǀ**M**

[100] Ibid.
[101] This example is from Thread 045 about American actor Jack Nicholson.

6.6 Chapter summary

In this chapter it has been established that DR elements are at their most prominent in two element subcategories: suspensive message-oriented (|m) or interactional (|oi) elements. These act as a type of preface for the upcoming linear unit. The salience of certain words in the DR elements of these two subcategories has also been established in this chapter. The following were explored: |m elements containing *say* and *saying*; and |oi containing *guess*, *mean*, *say* and *LOL*. It was found that linear units that contain these elements perform a variety of roles in the linear structure of the IMDb corpus, often depending on their turn orientation. Amongst the findings were that linear units containing these suspensive elements often act at the end of the turn as a means of signalling a reiteration of what has been said and as a means of signalling that the immediate linear connection is not the most relevant connection to make and that the reader should go back to the shared records of previous information from the thread to make best sense of what is coming up. It was also found that in linear units with inter-turn orientation, these DR suspensive elements appear in compliant linear units signalling conviviality by supporting the previous poster's turn. However, more often they occur in suspensive linear units and in such cases are better thought of as signals of antagonism and conflict. This will be explored further in Chapter 8.

Having explored elements in the IMDb corpus, firstly in non-structural terms in Chapter 5 and in terms of element relations and linear unit relations in this chapter, we now move

on in Chapter 7 to examining the relative frequencies of the categories and subcategories of linear units as well as the relative frequency of linear units containing DR elements.

CHAPTER 7

LINEAR UNITS AND DISCOURSE REFLEXIVITY IN THE IMDB CORPUS DISCOURSE

7.1 Overview

In Chapters 5 and 6, research questions 1–4 were addressed (see section 4.6). It was established that DR elements are more salient in certain element categories than others, namely in |oi and |m elements in initial position. These elements were then described in terms of the lexico-linear elements and sequences commonly present in each category and their role in the discourse. In this chapter, research questions 5, 6 and 7 will be examined by describing the relative frequencies and specific roles of linear units. This will be done through the description of:

1. The two main categories of linear units: message-oriented (M) and organizational-interactional (OI) linear units.

2. The two main categories when taking into consideration their linear compliance, i.e. whether they are core, compliant or suspensive linear units.

3. The two main categories when taking into consideration their turn orientation, i.e. whether those linear units which contain a prospection are oriented towards a completion within the same turn (pro intra-turn) or a completion in another turn (pro inter-turn) and whether those linear units which contain a

retrospective mechanism are oriented towards the same turn (retro intra-turn) or towards another turn (retro inter-turn).

4. Subcategories of M linear units: core (core M); compliant (M–, +M–, +M, MS and MS–); and suspensive (|M and |M–) linear units and the role of these subcategories in relation to their turn orientation.

5. The different types of retrospective supplementary mechanisms employed in compliant and suspensive M linear units.

6. Subcategories of OI linear units: compliant (OIS), and suspensive (|OI) linear units and the role of these subcategories in relation to their turn orientation.

7. The different types of retrospective supplementary mechanisms employed in compliant and suspensive OI linear units.

Having established the role of each category of linear units in general, each of these will then be re-examined comparing the frequency of linear units which contain DR elements compared to linear units which do not contain DR elements in order to find the points in the IMDb corpus discourse where significant differences occur in the behaviour of linear units that contain DR elements compared to those that do not.

7.2 Linear units in general: categorization, linear compliance and turn orientation

7.2.1 Linear units in general

Research Question 5: *How are linear units in general – and linear units that contain DR elements in particular – distributed in the IMDb corpus discourse?*

Research Sub-question 5a*: What are the relative frequencies of the two main categories of linear units: M and OI linear units?*

It can firstly be seen in Table 7.1 that there is a total of 4,159 linear units in the IMDb corpus. With a total number of words of 41,195 words and 15,102 elements (as seen in section 5.1) this means that a linear unit in the IMDb corpus is an average of 9.91 words and 3.63 elements long.

It can also be seen in Table 7.1 that the vast majority of linear units, as we might expect from the results we have seen already regarding elements in Chapter 5, are M linear units, accounting as they do for over 95 percent of all linear units.

Table 7.1: Linear units of the two main categories

	Freq.	**%**
M	3954	95.07
OI	205	4.93
Total	***4159***	**100**

As was stated in Chapter 5, it is difficult to make conclusions about such a figure given the fact that no comparable study has been carried out employing the present linear-based model. However, from the high percentage of M linear units seen in Table 7.1 it would seem to corroborate that the IMDb corpus discourse is overwhelmingly message-oriented, albeit with a small number of OI linear units also present.

7.2.2 Linear units that contain DR elements in general

Research Sub-question 5b: *Are linear units that contain DR elements distributed in the same way as linear units that do not contain DR elements or are there significant differences in their distribution?*

It can be seen in Table 7.2 that over 15% of all linear units contain at least one DR element, thus confirming that discourse reflexivity represents an important feature of the IMDb corpus discourse.

Table 7.2: Linear units that contain DR elements

	Freq.	**%**
Linear units that contain DR elements	627	**15.08**
Linear units with no DR elements	3532	84.92
Total	***4159***	***100***

However, the percentage of linear units that contain DR elements varies according to the main linear unit categories. As already established in Chapter 5, in order to find out whether the differences in the distribution of linear units that contain DR elements and linear units that do not contain DR elements are significant, the Log Likelihood Ratio (G^2) can be used.

Table 7.3: Discourse reflexivity in linear units in general

	Observed frequencies		**Expected frequencies**		
	DR	**non-DR**	**DR**	**non-DR**	**G^2**

M	**587**	3367	**596.09**	3357.91	**-0.16**
OI	40	165	30.91	174.09	**2.93**

As can be seen in Table 7.3, the number of M linear units that contain DR elements, at 587, is approximately the same as what would be expected if linear units which contain DR elements were distributed in the same way as linear units which do not contain DR elements. In this case the expected figure is around 596. The G^2 figure is consequently close to zero at -0.16. Similarly, there is no significant difference in the distribution of OI linear units that contain DR elements compared to what would be expected if they were distributed in the same way as linear units which do not contain DR elements. A G^2 figure of 2.93 indicates the figure is higher but not enough to indicate significance.

In general, then, the corpus data indicates that there is no evidence of over-representation of DR linear units in either category, at least when considering the category as a whole. In sections 7.3 and 7.4, we will examine if these general findings hold when subcategories of each are examined.

In the next section, the data in general will be examined in relation to linear compliance and this will then be related again to the two main categories of linear units.

7.2.3 Linear compliance in linear units in general

Research Sub-question 5c*: What are the relative frequencies of linear units in terms of their linear compliance, i.e. whether they are core linear units, compliant or suspensive?*

Table 7.4: Distribution of linear units by linear compliance

	Freq.	**%**
Core	460	11.06
Compliant	3129	75.23
Suspensive	570	13.71
Total	***4159***	***100***

It can be seen in Table 7.4 that over 75% of linear units enter into a compliant relation with either the previous linear units, the upcoming linear unit or both. Nevertheless, the fact that almost 14 percent of linear units are suspensive, i.e. that in some way they flout the expectation created by the previous discourse would seem to be, in itself, a salient feature of the IMDb corpus discourse. However, what this figure does not reveal is the retro-turn orientation of these suspensive linear units, i.e. whether it is a writer suspending the linear expectations created within his/her own turn or whether it is a writer suspending the linear expectations created in another participant's turn. This crucial aspect of suspension will be explored in sections 7.3.3 and 7.4.3.

It can also be seen that over 11% of linear units are core linear units. As discussed in section 6.5, these are often stand-alone linear units at the end of a turn. This will be explored in more detail in Chapter 8. The relatively high number of core M linear units is also accounted for by a number of the thread subject lines being stand-alone core M linear units.

Table 7.5: Main categories of linear units by linear compliance

	Core	%	Compliant	%	Suspensive	%	Total	Total %
M	460	100	3037	97.06	457	80.18	*3954*	95.07
OI	0	0	92	**2.94**	113	**19.82**	*205*	4.93
Total	***460***	***100***	***3129***	***100***	***570***	***100.00***	***4159***	***100***

Table 7.5 divides the general results seen in Table 7.4 into the two main categories: M and OI linear units. It can be seen that the distribution of suspensive linear units between M and OI linear units differs from the distribution seen in the figures in Table 7.1 regarding all linear units.

As can be seen in Table 7.5, almost 20% of suspensive linear units are OI linear units whereas only slightly fewer than three percent of compliant linear units are OI linear units. Suspensive OI linear units will be discussed in section 7.4. As noted in section 4.3.3.6, all OI linear units are considered to be qualified status linear units in the IMDb corpus discourse, so all core linear units are M linear units.

7.2.4 Linear compliance in linear units that contain DR elements

Research Sub-question 5d: *Are linear units that contain DR elements distributed in the same way as linear units that do not contain DR elements in terms of their linear compliance or are there significant differences in their distribution?*

Table 7.6: Discourse reflexivity and linear compliance in linear units in general

	Observed frequencies		Expected frequencies		
	DR	**non-DR**	**DR**	**non-DR**	**G^2**

Core	42	418	69.35	390.65	**-14.44****
Compliant	426	2703	471.72	2657.28	-5.36
Suspensive	**159**	411	**85.93**	484.07	**61.18****

As can be seen in Table 7.6, the most striking feature of the data in terms of linear compliance is that the observed frequency of suspensive linear units that contain DR elements is 159 whereas the expected frequency is around 86. With a G^2 figure of 61.18 indicting strong significance, this shows that suspensive linear units are made up of an unexpectedly high number of linear units that contain DR elements as compared to what would be expected if they were distributed in the same way as linear units which do not contain DR elements. Given that it has been established that the high frequency of suspension in general is a prominent feature of the IMDb corpus discourse, this further serves to indicate the importance of discourse reflexivity.

It can also be seen from Table 7.6 that there is evidence of an under-representation of core linear units that contain DR elements, with a G^2 figure of -14.44. This finding is not too surprising since DR linear units, by definition, make some sort of reference to the present discourse and so, more often than not, signal a clear relation between the present linear unit and the previous or upcoming discourse. There is no evidence that compliant linear units that contain DR elements are any more prominent than what would be expected if they were distributed in the same way as linear units which do not contain DR elements.

In the next section, the turn orientation of linear units in general will be described. It will then go on to examine the distribution of DR linear units in this light.

7.2.5 Linear compliance and turn orientation in linear units in general

Research Sub-question 5e*: What are the relative frequencies of linear units in terms of their turn orientation, i.e. whether they are pro intra or pro inter-turn and whether they are retro intra or retro inter-turn linear units?*

It can be seen in Table 7.7 that the total number of linear units in the data which display a retrospective relation with the previous discourse, at 3,195, is almost three times as high as the total number of linear units which display a prospective relation, at 1,198 linear units. In terms of percentages of the total number of linear units, almost 77% of linear units demonstrate a retrospective relation, whilst less than 29% of linear units demonstrate a prospective relation. It can also be seen in Table 7.7 that intra-turn relations, i.e. pro intra-turn added to retro intra-turn linear units are over three times more frequent (at 3,449 linear units in total) than inter-turn relations, i.e. the total of pro inter-turn and retro inter-turn linear units (at 944 linear units in total). Although this is, at the most simple level, a manifestation of the fact that turns tend to be longer than one single linear unit, it also provides us with empirical data as to what extent IMDb corpus discourse is oriented directly towards interaction between the participants' turns. The conclusion that can be drawn at this stage is that, in a linear analysis, the IMDb corpus discourse is clearly more frequently oriented to the establishment of relations within a turn rather than with another turn.

Table 7.7: Linear units by turn orientation

	Pro turn			**Retro turn**		
	Pro intra	**Pro inter**	***Pro Total***	**Retro intra**	**Retro inter**	***Retro Total***
Total	975	223	***1198***	2474	721	***3195***
% of linear units	23.44	**5.36**	***28.81***	59.49	**17.34**	***76.82***

In this connection, it is interesting to note that only a very small percentage of linear units, at 5.36%, actually prospect a response from another participant as seen in pro inter-turn relations linear units in Table 7.7. These, as explained in Chapter 4, are typically what has been termed elicitation moves (Sinclair and Coulthard 1975; Tsui 1994) or eliciting moves (Francis and Hunston 1992) in previous related models of spoken discourse. This compares to over three times that number, at over eighteen percent of linear units, which have a retro inter-turn orientation.

The nature of this type of turn orientation will be explored further in sections 7.3.3 and 7.4.3.

7.2.6 Linear compliance and turn orientation in linear units that contain DR elements

Research Sub-question 5f: *Are linear units that contain DR elements distributed in the same way as linear units that do not contain DR elements in terms of their turn orientation or are there significant differences in their distribution?*

The most salient result evident in Table 7.8 is that the observed frequency of retro inter-turn linear units that contain DR elements, at 167, is much higher than the expected frequency of around 118. As seen in the G^2 figure of 16.17, this difference is of strong significance. This indicates that linear units that contain DR elements are employed in the IMDb corpus discourse as a means of connecting retrospectively with another participant's turn much more frequently than what would be expected if linear units which contain DR elements were distributed in the same way as linear units which do not contain DR elements. In contrast, none of the other categories of linear units sorted by their turn orientation which contain DR elements are either over or under-represented to any significant level. This means that linear units that contain DR elements would seem to occur more or less as frequently as what would be expected if they did not contain DR elements in these categories.

Table 7.8: Discourse reflexivity and turn orientation in linear units in general

	Observed Frequencies		**Expected Frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Pro intra	169	806	131.28	799.72	6.91*
Pro inter	38	185	32.29	196.71	0.77
Retro intra	367	2107	348.01	2119.99	0.56
Retro inter	**167**	554	**118.59**	722.41	**16.17****

Whether all of the findings presented in this chapter thus far are applicable across both main linear unit categories (i.e. M and OI linear units) will be explored in sections 7.3 and 7.4 respectively.

7.3 M linear units: categorization, linear compliance and turn orientation

7.3.1 Linear compliance in M linear units in general

Research Question 6: *How are M linear units in general – and M linear units that contain DR elements in particular – distributed in the IMDb corpus discourse?*

Research Sub-question 6a*: What are the relative frequencies of the compliant subcategories of M linear units (M, M–, +M, +M–, MS, MS–) and suspensive subcategories of M linear units (\M and \M–)?*

Table 7.9 provides us with more precise information as compared to Table 7.5 regarding M linear units by breaking down the three linear compliance categories into the subcategories of M linear units. What can be seen, then, is that the majority of M linear units, at almost 77% are compliant M linear units. Within the subcategories of compliant M linear units, it can be seen that MS linear units is by far the most frequent category at about 32% of all M linear units.

From the descriptions of linear unit relations in Chapter 4, it can be concluded there are two basic types of compliant M linear unit sequences (see Coulthard et al (1981) for an alternative definition of the term *sequence* in a linear description of discourse). The first is an M– / +M sequence, i.e. an incomplete prospecting linear unit which is completed by the second linear unit. The second type of linear sequence is an M / MS sequence, whereby a linear unit, which is itself complete and so does not prospect completion, is

followed by a second linear unit displaying a retrospective supplementary mechanism.[102] [103] The total number of completed prospections (i.e. the total of +M and +M–) is 1,006, whereas the total number of compliant M linear units which contain a retrospective supplementary mechanism (i.e. MS and MS– linear units) is higher at 1,539. It can therefore be concluded that the M / MS sequence is the more frequent in the IMDb corpus discourse. Section 7.3.2 will explore in which of these two types of linear sequences linear units that contain DR elements play a more prominent role.

Table 7.9: M linear units by linear compliance

Core	**Total**	**% of M linear units**
Core	460	11.63
Total	***460***	***11.63***
Compliant	**Total**	**% of M linear units**
M–	492	12.44
+M	826	20.89
+M–	180	4.55
MS	1279	**32.35**
MS–	260	6.58
Total	***3037***	***76.81***
Suspensive	**Total**	**% of M linear units**
|M	287	7.26
|M–	170	4.30
Total	***457***	***11.56***
Total M linear units	***3954***	***100***

We can also see in Table 7.9 that suspensive M linear units represent over 11% of M linear units, with around two thirds of these, at 7.26%, of all linear units, being |M linear units, i.e. those which do not contain a prospection. It is therefore much less frequent for

[102] As can be seen in section 8.2.3, an alternative to the M / (MS) sequence is M / (OIS), when the second linear unit is more interaction-oriented than message-oriented. Here, however, the focus is exclusively on M linear unit sequences.

[103] It is possible that the first linear unit of these two sequence types displays retrospective relations to the previous linear unit. Hence, for example, the first linear unit of a compliant M– / +M sequence may be M–, +M– or MS–. For simplicity's sake, all of these will be referred to as being an M– / +M sequence.

a suspensive M linear unit to contain a prospection. The frequency of the turn orientation of these prospecting |M– linear units will be explored in section 7.3.3.

By definition, suspensive M linear units display some sort of retrospective supplementary mechanism. Hence, if we add the total number of suspensive M linear units to the total number of compliant MS and MS– linear units, we will reach the total number of M linear units which contain a retrospective supplementary mechanism. The total number, as seen in Table 7.10 is 1,996, which, in percentage terms, means that over 66% of all M linear sequences display some sort of retrospective supplementary mechanism. This compares to the total of 1,006 which represent a completed M– / +M sequences. Thus, it can be seen that retrospective supplementary mechanisms (in M / MS, M– / |M or M /|M sequences) are around twice as frequent in the IMDb corpus discourse as prospection/ completion (in M– / +M sequences).

Table 7.10: M linear units by linear sequence

	Total	**% of M sequences**
M– / +M sequence	1006	33.51
M / MS sequence[104]	**1996**	**66.49**
Total	***3002***	***100***

As well as providing empirical evidence for Sinclair's (1993/2004e) assertion that the presence of a retrospective supplementary mechanism in a linear unit can be conceived as being the default hypothesis, what the system of analysis employed here also reveals is to

[104] In this table as well as in Tables 7.14 and 7.17, the category of M/ MS sequence also includes M– / |M and M /|M sequences

what extent the retrospective supplementary mechanisms used by participants in IMDb discourse either add to the discourse in a compliant manner or, alternatively, to what extent they suspend expectations. As can be seen from Table 7.9, the frequency of compliant linear units containing a retrospective supplementary mechanism (total of MS and MS–), at 1,539, is much higher than the figure for suspensive linear units (total of |M and |M–), at 457. However, although suspensive M linear units are much less frequent than their compliant counterparts, they do represent over 11% of all M linear units and so would seem to represent an important characteristic of the IMDb corpus discourse, nonetheless.

In section 7.3.4 the question of whether these findings are universal or associated to one type of turn orientation will be explored.

7.3.2 Linear compliance in M linear units that contain DR elements

Research Sub-question 6b: *Are M linear units which contain DR elements distributed in the same way as M linear units which do not contain DR elements or are there significant differences in their distribution?*

Examining firstly compliant M linear units, we can see in Table 7.11 that, in general terms, the observed frequencies of the different subcategories of compliant linear units that contain DR elements are approximately what would be expected if they were distributed in the same way as linear units which do not contain DR elements. The exception to this is the subcategory of +M linear units whose observed frequency is

significantly lower than expected with a G^2 figure of -22.04. This indicates that M linear units which contain DR elements occur as a completion less than what would be expected if they were distributed in the same way as linear units which do not contain DR elements.

Table 7.11: Discourse reflexivity in compliant M linear units

	Observed frequencies		**Expected frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
M–	73	419	74.17	417.83	-0.02
+M	79	747	124.53	701.47	**-22.04****
+M–	14	166	27.14	152.86	-8.84*
MS	188	1091	192.82	1086.18	-0.14
MS–	51	209	39.20	220.80	3.89

In contrast to the above, the frequency of DR suspensive M linear units is significantly higher than would be expected if distributed in the same way as linear units that do not contain DR elements (see Table 7.12). This finding is applicable to both |M linear units and |M– linear units, as can be verified by the high G^2 figures of 42.44 and 27.83 respectively.

Table 7.12: Discourse reflexivity in suspensive M linear units

	Observed frequencies		**Expected frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
|M	87	200	43.27	243.73	**42.44****
|M–	53	117	25.63	144.37	**27.83****

As regards core M linear units we have already seen in Table 7.6 that there is some evidence of a lower figure than would be expected of DR M core linear units. At -14.44 the G^2 figure reported in Table 7.13 indicates that this is of strong significance.

Table 7.13: Discourse reflexivity in core M linear units

	Observed frequencies		**Expected frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Core M	42	418	69.35	390.65	**-14.44****

Table 7.14 categorizes DR and non-DR M linear units as those which represented a completed M– / +M linear sequence (a total of +M and +M–) and those which represent a retrospective supplementary mechanism (a total of MS, MS–, |M, |M–). It can be seen that the observed frequency of linear units that contain DR elements in M / MS linear sequences, at 379, is significantly higher than expected with a G^2 figure of 22.36. In contrast, the observed frequency in completion linear units of the M– / +M sequence at 93 linear units, compared to the expected frequency of over 141 linear units indicates an unexpectedly low observed frequency of linear units that contain DR elements with a similarly robust G^2 figure of -30.30.

Table 7.14: Discourse reflexivity in M linear units by linear sequence

	Observed frequencies		**Expected frequencies**		

	DR	Non-DR	DR	Non-DR	G^2
M– /+M sequence	93	913	151.66	854.34	**-30.30****
M / MS sequence	379	1617	300.91	1695.09	**22.36****

It can be concluded that DR M linear units occur more than would be expected if they were acting like linear units that do not contain DR elements in the second linear unit of an M / MS sequence. In contrast, they occur less than expected in the second linear unit of an M– / +M sequence.

How these findings apply when taking into account the turn orientation of M linear units will be addressed in the next section.

7.3.3 Linear compliance and turn orientation in M linear units in general

Research Sub-question 6c*: What are the relative frequencies of M linear units in terms of their turn orientation, i.e. whether they are pro intra or pro inter-turn or whether they are retro intra or retro inter-turn linear units?*

Table 7.15 shows the different subcategories of compliant M linear units and their frequency when taking into consideration their turn orientation, i.e. whether they are prospective or retrospective and whether they are oriented in their prospection and retrospection to another linear unit within the same turn (intra-turn orientation) or to a linear unit in another turn (inter-turn orientation). As will be remembered, Table 7.7 showed that retro-turn orientation is much more frequent than pro-turn orientation. It can now be seen in Table 7.15 that that assertion holds for compliant M linear units with only

around 23% of M linear units displaying a pro-turn orientation but over 64% displaying retro-turn orientation.

Table 7.15: Compliant M linear units by turn orientation

Pro-turn						
	Pro intra	**% M LUs**	**Pro inter**	**% M LUs**	**Pro Total**	**% M LUs**
M–	404	10.22	88	2.23	492	12.44
+M	~	~	~	~	~	~
+M–	165	4.17	15	0.38	180	4.55
MS	~	~	~	~	~	~
MS–	208	5.26	52	1.32	260	6.58
Total	***777***	***19.65***	***155***	***3.93***	***932***	***23.57***
Retro-turn						
	Retro intra	**% M LUs**	**Retro Inter**	**% M LUs**	**Retro Total**	**% M LUs**
M–	~	~	~	~	~	~
+M	715	18.08	111	2.81	826	20.89
+M–	165	4.17	15	0.38	180	4.55
MS	1193	30.17	86	2.18	1279	32.37
MS–	211	5.34	49	1.24	260	6.58
Total	***2284***	***57.79***	***261***	***6.60***	***2545***	***64.39***

Within those compliant linear units with a pro-turn orientation, pro inter-turn orientation (i.e. these linear units where a response from another participant is prospected) is much less frequent in total at 3.93% of all M linear units and in all three subcategories (M–, +M– and MS–) when compared to their pro intra-turn figure equivalent, which totals over 19% of all M linear units. As regards retro-turn compliant M linear units, the category of

retro intra-turn linear units, at a total of over 57% of all M linear units, is over eight times as frequent as retro inter-turn linear unit, at 6.6% of all M linear units, This large imbalance applies to all four subcategories.

Table 7.16 shows the same type of results but for suspensive M linear units. It can be seen that, just as for compliant M linear units, retro-turn suspensive M linear units, at 11.56%, are more frequent than pro-turn suspensive linear units are at 4.3% of all M linear units. This in itself is unsurprising since, by definition, all suspensive linear units are retro-turn, as they are flouting an expectation set up by a previous linear unit. In terms of pro-turn orientation, just as with compliant linear units, the number of pro inter-turn suspensive linear units is much smaller than pro intra-turn suspensive linear units. However, the biggest difference between suspensive and compliant M linear units in terms of their turn orientation is seen in retro-turn orientation, in the bottom half of Table 7.16. Here it can be seen that retro inter suspensive M linear units, at 7.79% of all M linear units, is almost twice as frequent as retro intra-turn suspensive M linear units at 3.77% of all M linear units. This provides evidence that the suspension of linear expectations, at least when M linear units are involved, occurs more often in IMDb corpus discourse on occasions when a participants flouts the expectations created by another participant. Occasions when a participant flouts the expectations created by himself/herself within the same turn are much less common (see section 4.3.3.5 for examples of these).

Table 7.16: Suspensive M linear units by turn orientation

Pro-turn						
	Pro intra	**% M LUs**	**Pro inter**	**% M LUs**	**Pro Total**	**% M LUs**
|M	~	~	~	~	~	~
|M–	107	2.71	63	1.59	*170*	4.30
Total	**107**	**2.71**	**63**	**1.59**	***170***	**4.30**
Retro-turn						
	Retro intra	**% M LUs**	**Retro inter**	**% M LUs**	**Retro Total**	**% M LUs**
|M	106	2.68	181	4.58	287	7.26
|M–	43	1.09	127	3.21	170	4.30
Total	*149*	***3.77***	***308***	***7.79***	***457***	***11.56***

Indeed, we can see that the total of the two retro inter-turn |M linear unit subcategories, at 308 linear units, is actually higher than the combined total of all compliant retro inter-turn M linear units, at 261, as seen in Table 7.15. The fact that participants are more likely to flout the linear expectation of a previous contribution, by either not completing a prospection or contradicting a proposition than comply with expectations, is one of the most salient features of the corpus data. In non-structural terms, this means that that over half of the occasions when a participant in the IMDb corpus discourse decides to react to a previous contribution s/he does so by either challenging the very validity of the previous participant's elicitation or by supplementing the previous participant's contribution by means of contradiction. How this defines the character of the IMDb corpus discourse will be explored further in Chapter 8.

We already saw in Table 7.10 that M / MS linear sequences are around twice as frequent than M– / +M linear sequences. It can now be seen in Table 7.17, how this breaks down when taking turn orientation into consideration.

Table 7.17: M linear units by linear sequence and turn orientation

	Retro intra	**% M LUs**	**Retro inter**	**% M LUs**	**Retro Total**	**% M LUs**
M– / +M sequence	880	22.26	126	**3.19**	*1006*	*25.44*
M / MS sequence	1553	39.28	443	***11.20***	*1996*	*50.48*

It can be seen, then, that the M / MS linear sequences, employing retrospective supplementary mechanisms are more frequent in both retro intra and retro inter-turn orientated linear units than M– / +M sequences. This is most pronounced in retro inter-turn oriented linear sequences, where a retrospective supplementary mechanism is employed at over 11% of all M linear units and is over three times as frequent as the prospection / completion sequence at just over 3% .

7.3.4 Linear compliance and turn orientation in M linear units that contain DR elements

Research Sub-question 6d: *Are M linear units which contain DR elements distributed in the same way as M linear units that do not contain DR elements in terms of their turn orientation or are there significant differences in their distribution?*

Table 7.18 shows the difference in distribution between DR and non-DR compliant M linear units as regards their pro-turn orientation. There is no evidence from the data that compliant DR elements are unexpectedly more frequent in any of the subcategories in either pro intra or pro inter-turn orientation.

Table 7.18: Discourse reflexivity in M linear units by pro-turn orientation

	Observed frequencies		**Expected frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Pro intra M–	65	339	60.91	343.09	0.32
Pro intra +M–	13	152	24.87	140.13	-7.86*
Pro intra MS–	43	165	31.36	176.64	4.65
Pro inter M–	10	78	13.27	74.73	-1.02
Pro inter +M–	1	14	2.26	12.74	-1.01
Pro inter MS–	10	42	7.84	44.16	0.65

In Table 7.19 we can see that in the majority of subcategories of retro-turn compliant M linear units there is no significant difference between the observed and expected frequencies. However, there is some indication of an under-representation, albeit of weak significance, in those linear units which provide a completion to a prospection, both within the turn (in retro intra-turn +M linear units and retro intra-turn +M– linear units) and between turns (in retro inter-turn +M linear units). This would seem to indicate that DR elements occur less than would be expected in completion M linear units with orientation both within and between turns.

Table 7.19: Discourse reflexivity in compliant M linear units with retro-turn orientation

	Observed frequencies		Expected frequencies		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Retro intra +M	78	637	107.79	607.21	**-10.56***
Retro intra +M–	12	153	24.87	140.13	**-9.40***
Retro intra MS	190	1004	180.00	1014.00	0.64
Retro intra MS–	43	168	31.81	179.19	4.26
Retro inter +M	7	104	16.73	94.27	**-8.24***
Retro inter +M–	3	12	2.26	12.74	0.26
Retro inter MS	11	75	12.97	73.03	-0.37
Retro inter MS–	10	39	7.39	41.61	1.00

The findings of Tables 7.18 and 7.19, indicating that there is no over-representation of M linear units containing DR elements in compliant M linear units, would seem to represent a clear contrast with the findings so far related to suspensive DR M linear units as seen in Table 7.12. In Table 7.20, we can see a refinement of the observation that DR suspensive M linear units are unexpectedly frequent. It is, in fact, pro intra-turn |M– linear units that contain DR elements which are over-represented, whereas pro inter-turn |M– linear units are not. Perhaps it is more accurate to say, given the small numbers, that the data is so limited that we cannot provide substantial enough evidence in this category. In any case, the data shows that linear units that contain DR elements are unexpectedly frequent on those occasions when the writer flouts the expectations of the previous linear unit and then continues within the same turn.

Table 7.20: Discourse reflexivity in suspensive M linear units with pro-turn orientation

	Observed frequencies		Expected frequencies		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Pro intra |M–	38	69	16.13	90.87	**27.13****
Pro inter |M–	15	48	9.50	53.50	3.29

Table 7.21 shows the distribution of DR suspensive M linear units in relation to their retro-turn orientation. It can be seen that both retro inter-turn subcategories show a very significant over-representation of DR M linear units (G^2 = 32.63 and 17.90) meaning that we can be confident in stating that M linear units that contain DR elements that flout the discourse expectations created by the previous linear unit between turns are more frequent than what would be expected if they were distributed in the same way as linear units which do not contain DR elements. There is also an over-representation of DR elements evident in M linear units that contain DR elements that flout the discourse expectations created by the previous linear unit within the same turn, albeit of weak significance (G^2 = 10.51 and 10.30).

Table 7.21: Discourse reflexivity in suspensive M linear units

	Observed frequencies		Expected frequencies		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Retro intra |M	29	77	15.98	90.02	**10.51***
Retro intra |M–	15	28	6.48	36.52	**10.30***
Retro inter |M	58	123	27.29	153.71	**32.63****
Retro inter |M–	38	89	19.15	107.85	**17.90****

Through these calculations we have able to build up an ever-more refined profile of the role of M linear units containing DR elements in the linear structure of IMDb corpus discourse. It has been seen that in almost every subcategory and turn orientation, M linear units containing DR elements appear as suspensive M linear units more than would be expected if they were simply acting like linear units that do not contain DR elements. They are particularly prominent in inter-turn categories of retro-oriented linear units and in |M– linear units which prospect a continuation in the same turn. This finding would seem to be consistent with what is proposed by Sinclair (1992:87), as seen in section 2.4.4, i.e. that 'the subject matter becomes the discourse itself' in what is called in this study suspensive retro inter-turn linear units.

In contrast, there would seem to be evidence that in compliant M linear units, M linear units which contain DR elements simply occur for the most part in the same way as linear units that do not contain DR elements, both in terms of pro and retro orientation. This does not mean, however, that compliant linear units that contain DR elements are infrequent. Indeed, the category of retro intra MS linear units remains the most frequent subcategory of all linear units that contain DR elements.

7.3.5 Retrospective supplementary mechanisms in M linear units

Research Sub-question 6e*: What are the relative frequencies of M linear units in terms of the retrospective supplementary mechanism present, i.e. whether an encapsulation or overlay is present in the linear unit?*

As described in Chapter 4, the type of retrospective supplementary mechanism seen in M / MS linear unit relations or in M / |M relations can belong to one of two types: encapsulation or overlay.

Table 7.22: Retrospective supplementary mechanisms in M linear units

	Total	**% of Retro M linear units**
Encapsulation	**1398**	**70.04**
Overlay	598	29.96
Total	***1996***	***100***

It can be seen in Table 7.22 in the IMDb corpus discourse that the most frequent of these mechanisms is encapsulation at 1,398, at over 70% of all retrospective M linear units, (i.e. MS, MS–, |M and |M– linear units). The category of encapsulation covers a wide range of linguistic features from explicit encapsulation, seen in labelling noun phrases, such as *this sentiment,* to more implicit encapsulations such as the presence of logical connectives at the beginning of the linear unit (see section 4.3.3.7). The category of overlay represents almost 30%.

7.3.6 Retrospective supplementary mechanisms in M linear units that contain DR elements

Research Sub-question 6f: *Are M linear units which contain DR elements distributed in the same way as M linear units that do not contain DR elements in terms of the retrospective supplementary mechanism present or are there significant differences in their distribution?*

As can be seen in Table 7.23, it is clear the linear units that contain DR elements feature most prominently in those linear units which contain an encapsulation, with 269 occurrences as compared to the expected figure of around 211. This produces a G^2 of 17.69. In contrast, the category of overlay does not show a significant difference between the expected and observed frequency, thus providing evidence that DR elements do not play a particularly prominent role in overlay.

Table 7.23: Discourse reflexivity in M linear units with retrospective supplementary mechanism

	Observed frequencies		**Expected frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Encapsulation	**269**	1129	210.76	1187.24	**17.69****
Overlay	77	521	90.15	507.85	-2.36

In order to refine this initial finding, we can further categorize encapsulation by examining the cases of encapsulation using the three levels of explicitness as outlined in section 4.3.3.7. In this way, it was hoped to find which of the three levels of explicitness of encapsulation in M linear units accounts for the strong presence of discourse reflexivity seen in Table 7.23. From Table 7.24, it is clear that it is, in fact, the categories of both explicit and implicit encapsulation in which there is a very prominent presence of linear units that contain DR elements. Explicit encapsulation, it will be remembered are those cases which contain a determiner + labelling noun sequence, whereas implicit encapsulation covers a wide range of features from the presence of o elements in initial

position, verbs and nouns where the reader has to infer what is being encapsulated and so on. The fact that discourse reflexivity is less prominent in fuzzy encapsulation may be to be expected as encapsulation of this type is achieved through the use of demonstrative pro-forms, which are not considered to be discourse reflexive in this study.

Table 7.24: M linear units by explicitness of encapsulation

	Observed frequencies		**Expected frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Explicit encapsulation	34	49	12.51	70.49	**32.34****
Fuzzy encapsulation	63	232	44.47	250.53	8.23*
Implicit encapsulation	224	796	153.77	866.23	**33.92****

Following the findings so far in this chapter, one would assume that this strong presence of linear units that contain DR elements in encapsulation would be most prominent in suspensive linear units. We can see in Table 7.25 that this assertion is accurate, at least for M linear units containing explicit and implicit encapsulation with DR elements, as can be seen in the extremely high G^2 figures of 40.76 and 65.06 respectively. There would also seem to be a general over-representation of encapsulation in compliant M linear units but this tendency is not strong enough to be deemed significant.

Table 7.25: Discourse reflexivity in M linear units by types of encapsulation and linear compliance

	Observed frequencies		**Expected frequencies**		
Types of encapsulation	**DR**	**Non-DR**	**DR**	**Non-DR**	**G^2**
Compliant explicit	16	42	8.74	49.26	5.95

Compliant fuzzy	47	177	33.77	190.23	5.56
Compliant implicit	154	693	127.69	719.31	6.06
Suspensive explicit	18	7	3.77	21.23	**40.76****
Suspensive fuzzy	16	55	10.70	60.30	2.75
Suspensive implicit	70	103	26.08	146.92	**65.06****

It can therefore be concluded that there is strong tendency for DR elements to feature more than would be expected in suspensive M linear units containing an encapsulation and DR elements.

The encapsulation and overlay mechanisms can also be examined in terms of their linear compliance and turn orientation. Table 7.26 shows that DR elements are at their most prominent in suspensive M linear units, especially with inter-turn orientation, with a G^2 figure of 62.78. It can also be seen that there may be evidence that DR elements are prominent in suspensive M elements with intra-turn orientation containing an encapsulation although this is with weak significance.

As regards overlay, in general the finding above that suggests that there was no over or under-representation of DR elements in M linear units containing overlay was found to be upheld when taking into account linear compliance and turn orientation. An exception may be the category of suspensive inter-turn linear units where there is some evidence, albeit with weak significance ($G^2 = 7.42$), that DR elements are over-represented in these cases.

Table 7.26: Discourse reflexivity in M linear units containing encapsulation and overlay by linear compliance and turn orientation

	Observed frequencies		Expected frequencies		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Compliant retro intra-turn encap.	188	878	160.71	905.29	5.22
Compliant retro inter-turn encap.	15	58	11.01	61.99	1.56
Suspensive retro intra-turn encap.	32	88	18.09	101.91	**10.67***
Suspensive retro inter-turn encap.	60	79	20.96	118.04	**62.78****
Compliant retro intra-turn overlay	36	301	50.81	286.19	-5.56
Compliant retro inter-turn overlay	6	64	10.55	59.45	-2.67
Suspensive retro intra-turn overlay	5	22	4.07	22.93	0.24
Suspensive retro inter-turn overlay	38	126	24.72	139.28	**7.42***

7.4 OI linear units: categorization, linear compliance and turn orientation

7.4.1 Linear compliance in OI linear units in general

Research Question 7: *How are OI linear units in general – and OI linear units that contain a DR element in particular – distributed in the IMDb corpus discourse?*

Research Sub-question 7a*: What are the relative frequencies of OI linear units in terms of their linear compliance, i.e. whether they are compliant or suspensive linear units?*

As described in Chapter 4, there are two types of OI linear units: compliant OIS linear units; and suspensive |OI linear units in the IMDb corpus discourse.

Table 7.27: OI linear units by linear compliance[105]

	Freq.	***%***
OIS	92	***44.88***
|OI	113	***55.12***
Total	***205***	***100***

We can see in Table 7.27, then, that compliant OIS linear units represent almost 45% of OI linear units whereas there are 113 suspensive |OI linear units representing over 55% of OI linear units. As we have seen in section 7.2.2, such a figure relating to linear compliance is important in itself but does not provide the critical information as to whether these OI linear units are inter or intra-turn oriented. This will be explored in section 7.4.3.

7.4.2 Linear compliance in OI linear units that contain a DR element

Research Sub-question 7b: *Are OI linear units that contain a DR element distributed in the same way as OI linear units that do not contain a DR element in terms of their linear compliance or are there significant differences in their distribution?*

It should be noted again that the number of OI linear units that contain a DR element is, in terms of qualitative studies, relatively low, so any conclusions made in these sections must remain as suggestive rather than conclusive. From the results in Table 7.28, we can see that both categories would seem to have a slightly higher number of DR elements

[105] The relative distributions of the two categories relevant to this section are repeated from Table 7.5 in Table 7.27 for the reader's convenience.

present as compared to the expected frequency but neither of the related G^2 figures is significant.

Table 7.28 Discourse reflexivity in OI linear units by linear compliance

	Observed frequencies		**Expected frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
OIS	21	71	13.87	78.13	3.83
IOI	19	94	17.04	95.96	0.26

7.4.3 Linear compliance and turn orientation in OI linear units in general

Research Sub-question 7c: *What are the relative frequencies of OI linear units in terms of their turn orientation, i.e. whether they are pro intra or pro inter-turn or whether they are retro intra or retro inter-turn linear units?*

Table 7.29 shows sub-categories of OI linear units in relation to their turn orientation. From the data presented it can be seen that over 44% of IOI linear units prospect a continuation within the same turn. These are generally citations at the beginning of the turn using the [quote] [/quote] markup as well as a number of salutations at the beginning of a turn. It can also be seen that IOI linear units are very rarely pro inter-turn linear units, representing less than 3% of all OI linear units meaning that very few of them act as interactive elicitations such as *What?*. In terms of retro-orientation, the most notable feature is that IOI linear units with inter-turn orientation are almost twice as frequent as OIS linear units. These, on the whole, comprise citations using the [quote] [/quote] markup. In addition, over half of OIS linear units have retro inter-turn orientation,

meaning that they provide supportive acknowledgments for what has been said in the previous post.

Table 7.29 OI linear units by turn orientation

Pro turn			
	Pro intra	**Pro inter**	***Pro Total***
IOI	91	5	*96*
OIS	~	~	~
Total	91	5	*96*
% of all OI linear units	***44.39***	***2.44***	
Retro turn			
	Retro intra	**Retro inter**	***Retro Total***
IOI	4	96	*100*
OIS	40	52	*92*
Total	44	148	*192*
% of all OI linear units	***21.46***	***72.20***	

7.4.4 Linear compliance and turn orientation in OI linear units that contain a DR element

Research Sub-question 7d: *Are OI linear units that contain a DR element distributed in the same way as OI linear units that do not contain a DR element in terms of their turn orientation or are there significant differences in their distribution?*

The results in Table 7.30 would seem to indicate that the only category in which OI linear units is unexpectedly high with high significance is that of compliant retro inter-turn OIS linear units with DR elements with a G^2 figure of 19.45. The fact that it is compliant OIS

linear units which are over-represented and not suspensive |OI linear units is in definite contrast to the opposite situation with M linear units with inter-turn orientation.

Table 7.30: Discourse reflexivity in OI linear units by linear compliance and turn orientation

	Observed frequencies		**Expected frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
OIS retro intra	0	40	6.03	33.97	-13.07**
OIS retro inter	21	31	7.84	44.16	**19.45****
|OI retro intra	1	3	0.60	3.40	0.27
|OI retro inter	18	78	14.47	81.53	0.95
|OI retro null	0	13	1.96	11.04	-4.25
|OI pro intra	10	81	13.72	77.28	-1.29
|OI pro inter	2	3	0.75	4.25	1.82
|OI pro null	7	10	2.56	14.44	6.72

7.4.5 Retrospective supplementary mechanisms in OI linear units

Research Sub-question 7e: *What are the relative frequencies of OI linear units in terms of the retrospective supplementary mechanism present i.e. whether an encapsulation or verbal echo is present in the linear unit?*

As explained in Chapter 4, there are two types of retrospective supplementary mechanism which may be present in OI linear units: encapsulation and verbal echo. It can be seen in Table 7.31 that verbal echo constitutes over 43% of all retrospective OI linear units whereas over 56% of retrospective OI linear units contain an encapsulation.

Table 7.31: OI linear units by type of retrospective supplementary mechanism

	Freq.	**% of Total**
Encapsulation	109	**56.77**
Verbal Echo	83	**43.23**
Total	***192***	***100***

As noted in Chapter 4, one of the distinguishing characteristics of OI linear units is the fact that they often encapsulate the previous discourse using what can be characterized as an inexplicit encapsulation such as *ah* or *uhm*. It can be seen in Table 7.32 that such implicit encapsulation accounts for the vast majority of OI linear units at almost 92% of OI linear units with a retrospective orientation.

Table 7.32: OI linear units by type of encapsulation

	Freq	**% of Total**
Explicit encap.	0	0
Fuzz encap.	9	8.26
Implicit encap.	100	91.74
Total	***109***	***100***

7.4.6 Retrospective supplementary mechanisms in OI linear units that contain a DR element

Research Sub-question 7f: *Are OI linear units that contain a DR element distributed in the same way as linear units that do not contain DR elements in terms of their*

retrospective supplementary mechanism or are there significant differences in their distribution?

Following from the findings in Table 7.32, it can be seen in Table 7.33 that the majority of OI linear units which contain a DR element actually represent the category of implicit encapsulation. This category has a significantly higher number of OI linear units containing an implicit encapsulation and a DR element than would be expected than if they did not contain a DR element.

Table 7.33: Discourse reflexivity in OI linear units by type of encapsulation

	Observed frequencies		**Expected frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Fuzzy encap.	1	8	1.36	7.64	0.12
Implicit encap.	29	71	15.08	84.92	**12.51****
Verbal echo	10	73	12.51	70.49	0.63

It can also be seen that there is no evidence of there being an unexpectedly high frequency of DR elements in OI linear units containing a verbal echo. This is perhaps unsurprising as these linear units comprise those parts of text quoted, usually using the [quote] [/quote] markup. In such cases one might expect a random selection of DR elements to be included.

If we look at OI linear units which contain an encapsulation in relation to their retro-turn orientation, we can see in Table 7.34 that the observed frequencies in OI linear units with

retro inter-turn orientation containing an encapsulation and a DR element is significantly more frequent than would be expected ($G^2 = 32.15$). This would seem to indicate that OI linear units with an encapsulating DR element are predominantly used for inter-turn commentary and evaluation.

Table 7.34 Discourse reflexivity in encapsulating OI linear units

	Observed frequencies		**Expected frequencies**		
	DR	**Non-DR**	**DR**	**Non-DR**	**G^2**
Encap. retro intra	1	43	6.63	37.37	**-8.29**
Encap. retro inter	29	36	9.80	55.20	**32.15****

7.5 Chapter summary

It has been shown in this chapter that the vast majority of linear units in the IMDb corpus discourse are M linear units with only a small percentage being OI linear units. It is also found that the number of linear units which show a retrospective orientation is over three times as high as those that show a prospective orientation indicating that the IMDb corpus is for the most part retrospective in its orientation. It is particularly infrequent for linear units to prospect a contribution from another poster.

Within M linear units, there is an important number which are suspensive. It was also found that retro inter-turn suspensive M linear units are almost three times as frequent as retro intra-turn suspensive M linear units, meaning that the suspension of linear expectations in M linear units occurs more often on occasions when a participants flouts the expectations created by another participant than within the same turn. It was

concluded that in over a half of the occasions when a participant in the IMDb discourse responds to a previous turn s/he is challenging the validity of the previous participant's elicitation or contradicting it in some way.

It was also found that in the IMDb corpus that the M / MS sequence is around twice as frequent as the M– / +M sequence meaning that the retrospective supplementary mechanism of encapsulation and overlay are around twice as frequent as prospection/completion. The most frequent retrospective supplementary mechanism present in M / MS sequences is encapsulation.

In OI linear units, slightly over half are suspensive |OI linear units. Almost half prospect a continuation. These are generally verbal echoes. Verbal echoes constitute almost 38% of all OI linear units whereas over 67% of OI linear units contain an encapsulation.

It has also been shown that discourse reflexivity is an important aspect of IMDb corpus discourse, with over 12% of linear units featuring at least one DR element. Discourse reflexivity was found to be salient in certain subcategories of both M and OI linear units depending on their linear compliance and turn orientation. Most notably, it was found that discourse reflexivity is prominent in suspensive M linear units, both retro inter and retro intra-turn orientation, but especially retro inter-turn orientation, i.e. when contradicting or questioning the validity of the contribution of another participant. It was also shown that discourse reflexivity is prominent in the second part of M / MS sequences and is not prominent in M– /+M sequences. It is also prominent in inter-turn

OIS linear units. In those categories in which discourse reflexivity is prominent, it is in those linear units which contain an encapsulation, especially implicit encapsulation, that discourse reflexivity is at its most prominent.

As the suspensive linear units with retro inter-turn orientation containing an encapsulation (both |M– and |M linear units) are the most salient in terms of having the highest G^2 figures, these will be the focus of a qualitative discussion of an extended extract from a thread which constitutes Chapter 8.

CHAPTER 8

DISCOURSE REFLEXIVITY IN SUSPENSIVE RETRO INTER-TURN M LINEAR UNITS

8.1 Overview

In Chapters 5 and 6 findings of a quantitative analysis of the IMDb corpus regarding the frequency and salience of both elements in general and discourse reflexive (DR) elements in particular were presented. These findings were illustrated with examples of the most salient DR elements and their immediate surrounding context. In Chapter 7, qualitative findings as regards linear units in general and linear units containing DR elements in particular were presented. The emphasis thus far has therefore been on establishing the quantitative findings related to the system of analysis. However, it is also necessary to examine how these observations regarding frequency and significance are grounded in what happens in individual threads in the IMDb corpus discourse. Hence, this chapter will illustrate and discuss the nature of discourse reflexivity in linear units in the IMDb corpus discourse through a qualitative-based approach with particular reference to several extracts of threads from the IMDb corpus.

In Chapter 7, then, it was found that linear units containing DR elements are particularly salient in suspensive retro inter-turn M linear units, i.e. those which are oriented towards the previous turn. By definition, linear units in this category occur at the beginning of a turn acting as a response to a previous post and include moments where the poster 'challenges' (Burton 1980; Sinclair 1992; Warren 2006) what has been written by the

previous poster in some way (see section 2.4.4). It was also demonstrated that discourse reflexivity is most salient in those suspensive retro inter-turn M linear units in which an encapsulation is present, particularly explicit and implicit encapsulation. It was also found that discourse reflexivity is prominent, though to a lesser extent, in suspensive retro inter-turn M linear units which contain an overlay. In non-technical terms, this means that the posters in the IMDb corpus tend to refer to a prior stretch of the present discourse explicitly in moments when the expectations created by the previous linear unit have not been met between turns. In the case of encapsulation this is done through the labelling of a stretch of the previous discourse, whilst in the case of overlay it is done through the use and modification of a stretch of the previous discourse.

As will be seen below, such suspensive retro inter-turn M linear units containing both an encapsulation and DR elements almost exclusively provide some sort of negative evaluation of the previous turn and would seem to be one of the principal mechanisms by which the type of antagonism, as reported as being prevalent in similar forms of Internet communication (e.g. Pihlaja 2011; Arendholz 2013), occurs. It will also be seen that suspensive retro inter-turn M linear units containing both an overlay and DR elements are more varied in their function, serving as a means of antagonism but also as a means of clarification.

In order to explore these issues, a detailed analysis of an extended extract from the IMDb corpus discourse will first be presented in section 8.2. How discourse reflexivity

functions in the extended extract seen in section 8.2 will then be explored as well as in further examples from the IMDb corpus in section 8.3.

8.2 The *Alien* thread

8.2.1 Introduction

In this section, issues related to discourse reflexivity will initially be set aside in order to focus on the linear structure of an extended extract taken from Thread 059 from the IMDb message board dedicated to the 1979 science fiction film, *Alien* (hereafter the *Alien* thread). The *Alien* thread was chosen as it represents a thread containing a number of suspensive retro inter-turn M linear units including both encapsulation and overlay in several different M / |M sequences. It is also a thread in which DR elements play a prominent role in these retrospective supplementary mechanisms.

Firstly, then, the *Alien* thread (see Example 8.1a) will be described in non-technical terms in section 8.2.2. This non-technical description will then be related to the linear turn structure of the extract in section 8.2.3 and to the retrospective supplementary mechanisms involved at the points where one turn connects retrospectively to the previous turn in section 8.2.4. Occurrences of DR elements in the extract will finally be examined in section 8.2.5.

8.2.2 Summary of the *Alien* thread

The extract (see Example 8.1a) is taken from the IMDb message board dedicated to the film *Alien*. In message boards dedicated to a particular film or person from the film

industry, regular posters tend to be established fans of the film or person. In consequence, a partisan tone is often prevalent and any criticism of the film or person is often met with a rather defensive reaction from such posters. The thread from which the extract is taken is one such thread.

Example 8.1a[106] [107]

P	T	L	Linear unit[108]
A	1	1.	This is the most overated sci fi of all time!
		2.	I cant believe so many people liked this movie!
		3.	It is so boring!
		4.	The acting was horrible
		5.	(he's a f*#%ing robot!")
		6.	and the effects are extreemly low quality.
		7.	It was obviously a low budget film
		8.	and it deserves none of the credit it is getting .
		9.	I rented this movie expecting a blockbuster
		10.	and I get a cheezy made for tv callibur snoozefest .
		11.	What is so good about this film?
		12.	This belongs with battlefeild earth and troll 2.
B	2	13.	um this was the 70s,
		14.	nothin super high-tech,
		15.	i mean, yeah, its not like its using KingKong cgi...
		16.	i think u need to re-think this.
		17.	NOTHING belongs with troll 2. NOTHING.
		18.	it didt scare me,

[106] In this chapter, Examples 8.1 and 8.2 are variously repeated entirely or selectively. These are given a letter for coding, e.g. 8.1a, 8.1b etc.
[107] As in previous chapters, the first three columns in these tables provide information on the Poster (P); Turn (T) and Line number (L).
[108] As in previous chapters, in linear unit example tables, double horizontal line indicates the end of a turn; a solid horizontal line indicates the end of a subject line; a dotted line indicates the end of a linear unit. A wavy line means that part of the same turn is omitted. A double wavy line means that part of discourse is missing from another turn. The posts are divided into linear units, one per line.

		19.	but i enjoy it alot
		20.	if u hate alien, watch AvP,
		21.	invite ur friends over, pop some popcorn and make fun of the entire movie.
		22.	sorry tha u didnt enjoy alien,
		23.	me, myself, i prefer Aliens
B	3	24.	o and when i said, i didnt scare me but i enjoyed it, i ment Alien, not troll 2,
		25.	tha movie sucks
C	4	26.	Wow, a kid "discussing" movies.
		27.	Incredible!
A	5	28.	Just because I dont agree with you does not mean you have to resort to name calling.
D	6	29.	He must be one of those AVP fanboys.
E	7	30.	More like one of those trollboys.
		31.	If you click on his screen name, you'll find similar posts by him on other boards.
A	8	32.	How am I a troll?
		33.	Oh I see, I have an opinion that you dont agree with and that I feel strongly about.
F	9	34.	Shut the f uck up <name A.>,
		35.	you're out of your element.
		36.	I did not watch my buddies die FACE DOWN IN THE MUCK
		37.	so this f ucking muppet, this f ucking creep could critizize one of the best films ever made...
		38.	the movie whose structural perfection is matched only by its sheer awesomeness.
G	10	39.	If I had a dime for every time someone called any movie "the worst of all time!" I'd be a millionare by now.

In this context, Poster A begins with what can be described as a provocative subject line for the thread in line 1, providing as it does, a rather hyperbolic negative evaluation of the film *Alien*. The poster continues in a similar vein throughout the body of the initial turn

(lines 2–12) restating the stance in a variety of ways (e.g. lines 2, 10, 11 and 12). The poster does not pose a genuine question to which other posters can respond, and instead uses, what can be called, in non-technical terms, a rhetorical question in line 11, i.e. a question to which the poster does not genuinely expect a response. The fact that Poster A states his opinion rather than asking for opinions may, it will be argued below, influence the subsequent negative reaction of the other posters.

Four posters (Posters B, C, F and G) respond directly to the original post (beginning at lines 13, 26, 34 and 39 respectively). Each of these expresses varying degrees of negativity towards Turn 1. These range from a courteous disagreement of the viewpoint expressed by Poster A in Turn 2 from Poster B, more ironic criticisms of the original posts in Turn 4 from Poster C and Turn 10 from Poster G to a more aggressive and seemingly insulting response in Turn 9 from Poster F. In these responses the subject matter also shifts from an exchange about the film itself in Turns 1 and 2 to an evaluation of the original post and the posters' behaviour in the other turns.

The extract also includes five subsequent turns which follow on from two of these responses. The first (Turn 3 from Poster B in line 24) is rather unusual in that it represents Poster B responding to him/herself in order to clarify a point s/he made in Turn 2. The second of these, Turn 5 in line 28, is a response by Poster A, to Turn 4. In Turn 5, Poster A criticizes the response provided by Poster C in Turn 4 and accuses Poster C of *name calling*. Thus, the appropriateness of Poster C's behaviour in the discourse becomes the focus of attention at this point. A second response to the Turn 4 is

seen in Post 6 by Poster D (line 29). However, in this case, Poster D aligns himself/herself with Post C by insinuating that Poster A has bad taste in films. This is seen in the derogatory term *AVP fanboy* (*AVP* being *Alien vs. Predator* a related film generally regarded as a sequel of very poor quality). Poster E (in line 30) then aligns him/herself with Poster D's criticism of Poster A. However, instead of calling Poster A an *AVP fanboy,* Poster E adapts the wording of Poster D's turn in order to coin the derogatory term *trollboy* to describe Poster A. In doing so, s/he makes a reference to intertextual links which s/he claims provides evidence of similar inappropriate behaviour of which Poster A has been guilty in other IMDb message boards. Poster A then shows indignation towards this accusation in Turn 8 by challenging it through a question *How am I a troll?*, which s/he then answers him/herself in the same turn in order to dismiss the accusation.

The thread continues in a similar vein after the end of the extract presented here for a further ten turns with several posters criticizing the original post and Poster A's behaviour. Poster A responds variously to these. It is a thread, then, which quickly moves away from the topic of the original post and what is supposedly to be discussed in the message board to a rather antagonistic exchange about the appropriateness or otherwise of the posters' behaviour within the discourse. Ironically, the fact that the other posters shift the focus of the discussion away from the film *Alien* to Poster A's supposedly inappropriate behaviour means that the thread is an example of what has been termed *thread drift* (Arendholz 2013:49), itself considered a type of inappropriate behaviour on such message boards.

It is essentially, then, a dispute between Poster A on one side and various other posters on the other. Such *metadiscussions*, where participants discuss the rules of communication and the acceptability of certain behaviour have been a reported feature of message board interaction since early research into Internet communication (e.g. Crystal 2001). In the case of the IMDb corpus, such an extended stretch of discourse in which the discourse itself is the topic is unusual, although not unique.

8.2.3 The turn structure of the *Alien* thread

The summary of the thread presented in section 8.2.2 can now be related to the linear analysis of the present study. As stated in section 7.3, one of the fundamental claims made in this study is that there are two basic compliant linear unit sequences in discourse:

M / (MS) or (OIS)

M– / +M

In the context of retro inter-turn linear units, the second linear unit in these sequences represents a response provided by a second writer. In the M / (MS) or (OIS) sequence, the initial M linear unit does not prospect that a linear unit will follow. However, it can be supplemented and supported by an MS linear unit if the second linear unit is message-oriented and by an OIS linear unit if the second linear unit is interaction-oriented. In an M– / +M sequence, on the other hand, the initial M– linear unit does prospect forward and thus requires a +M linear unit to complete the sequence.

One of the most notable characteristics of IMDb discourse, based on the findings in Chapter 7, was found to be the high frequency of occasions when expectations of

discourse are flouted by suspensive M linear units, particularly between turns. In inter-turn relations, if the second poster decides not to comply with the discourse expectations in place after the first linear unit and instead uses a suspensive linear unit, the equivalent sequences to M / (MS) or (OIS) and M– / +M sequences are: M / |M or |OI and M– / |M or |OI respectively[109]. As established in section 4.3.3.5, these can be described as follows:

1. M / |M or |OI: The initial statement provided by the first poster is challenged in terms of either it not reaching the target communicative state satisfactorily; its veracity; its appropriateness; or, in more extreme cases, the actual right the poster has to make such a statement.
2. M– / |M or |OI: The initial elicitation provided by the writer is not responded to as expected and instead the elicitation is challenged in terms of either it not communicating sufficiently in order for the second poster to respond successfully; or the poster's actual right to ask such a question.

As we have seen in Chapter 7, the first of these sequence types is much more frequent in the IMDb corpus discourse and for that reason it will form the bulk of the discussion below. In this connection, the *Alien* thread has been chosen based on the fact that inter-turn M / |M or |OI sequences containing an encapsulation or overlay as well as discourse reflexivity feature prominently.

[109] For clarity's sake, in general discussions the symbols M / |M will be used to refer to any sequence in which the first linear unit of the sequence is complete (i.e. M; +M; or MS) and the second is suspensive (i.e. |M or |M–). Similarly, M– / |M will be used in general discussions to refer to any sequence in which the first linear unit of the sequence is incomplete (i.e. M–; +M–; or MS–) and the second is suspensive (i.e. |M or |M–). In the detailed analysis of an individual sequence, the specific categories of that sequence will be referred to.

The nest turn structure of the extract of the thread can be seen in Figure 8.1. It is polylogic in nature (see section 5.4.1). It constitutes an initial turn (Turn 1) which is characterized as not containing an elicitation, i.e. that M– pro inter-turn linear units are not present in the turn, and, in consequence, it is a turn characterized as being a series of intra-turn M linear units.

Despite this, there are four direct responses to this initial turn, each of which begins with a suspensive M linear unit. As the initial turn does not included an M– pro inter-turn linear unit, the four direct responses to this initial turn (Turns 2, 4, 9 and 10) form part of an inter-turn M / |M sequence. Of these four responses, two turns (Turns 2 and 4) are responded to by subsequent turns. Turn 2 is responded to by the same poster in Turn 3 providing a self-clarification through a suspensive M linear unit and therefore forms an inter-turn M / |M sequence. Turn 4 is responded to by two separate turns. Turn 5 represents a challenge to Turn 4 by the Original Poster (A) through a suspensive M linear unit, and therefore again forms an inter-turn M / |M sequence. In contrast, Turn 6 represents a compliant MS linear unit, which supports Turn 4 and therefore forms an inter-turn M / MS sequence. Turn 6 is supported further by another MS linear unit in line 30 by Poster E again forming a compliant inter-turn M / MS sequence. At this point, Post 7 is challenged again by Poster A in the form of a suspensive M– linear unit in turn 8 to create a final inter-turn M/ |M sequence.

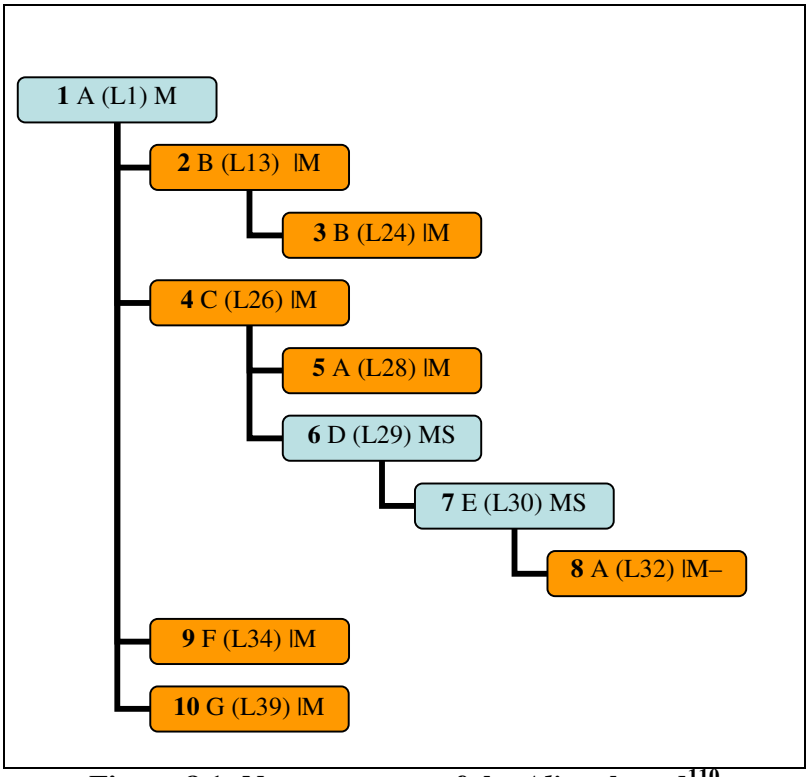


Figure 8.1: Nest structure of the *Alien* thread[110]

In terms of turn structure, then, we can see it is a thread characterized by a series of posters challenging the original poster's turn through the use of initial suspensive M linear units. Two posters then support these challenges through MS linear units. The original poster subsequently intervenes to challenge these challenges. Indeed, seven of the nine turns that represent a response in the extract begin with a suspensive M linear unit. The turn structure, therefore, is one which reflects the dispute as described in section 8.2.2 and so represents an embodiment of one of the findings presented in section 7.2.3, i.e. that suspensive inter-turn M linear units are more frequent than compliant inter-turn M linear units in the IMDb corpus.

[110] In this diagram each lozenge gives the following information: turn number; poster; line number; the first linear unit of turn (except the first turn in which the salient linear unit of the point of connection is given. Lighter shaded lozenge signifies compliant first linear unit; darker shaded lozenge suspensive first linear unit.

8.2.4 Points of connection and retrospective supplementary mechanisms

It is possible to gain further insights into the linear structure of the extract by looking at the point of connection of each sequence. As explained in section 4.3.3.8, the point of connection in an inter-turn sequence is the exact linear unit in the previous turn from which the second linear unit follows. This may or may not be the last linear unit of the previous turn in the nest structure of the thread. Instead, it may be any linear unit from the previous turn.

Table 8.1 shows the nine points of connection in the inter-turn M sequences from the extract in isolation. The second linear unit (in columns 3 and 4) forms a sequence with a linear unit in a previous turn (in columns 1 and 2). We can see that in some inter-turn M sequences the second linear unit of the sequence follows directly after the first. For instance, the fifth example in the table can clearly be interpreted as a linear sequence with the linear unit in line 30 in Turn 7: *More like one of those trollboys*, following immediately after the linear unit in line 29 in Turn 6: *He must be one of those AVP fanboys.* However, in other cases, the second linear unit of the sequence makes reference to a point much earlier in the discourse. For instance, in the final row in Table 8.1, the linear unit in line 39 in turn 10: *If I had a dime for every time someone called any movie "the worst of all time!" I'd be a millionare by now* follows on from the first line of the thread, i.e. the subject line, in line 1: *This is the most overated sci fi of all time!* and not line 12, the last linear unit of the first turn.

Table 8.1 Points of connection in inter-turn linear sequences

Turn: line	First linear unit of sequence	Turn: line	Second linear unit of sequence	Supp. Retro. Mech.
1: 6	*the effects are extreemly low quality.*	2: 13	*this was the 70s,*	Encap. (fuzzy)
2: 18–19	*it didt scare me but i enjoy it alot*	3: 24	*o and when i said, i didnt scare me but i enjoyed it, i ment Alien, not troll 2,*	Overlay
1: 12	*This belongs with battlefeild earth and troll 2.*	4: 26	*Wow, a kid "discussing" movies.*	Encap.(implicit)
4: 26	*a kid "discussing" movies.*	5: 28	*Just because I dont agree with you does not mean you have to resort to name calling*	Encap.(implicit)
6: 29	*He must be one of those AVP fanboys.*	7: 30	*More like one of those trollboys.*	Overlay
7: 30	*More like one of those trollboys.*	8: 32	*How am I a troll?*	Overlay
1:12	*This belongs with battlefeild earth and troll 2.*	9: 34	*Shut the f uck up <name A.>,*	Encap.(implicit)
1: 1	*This is the most overated sci fi of all time!*	10: 39	*If I had a dime for every time someone called any movie "the worst of all time!" I'd be a millionare by now.*	Overlay

The type of retrospective supplementary mechanism present in the second linear unit of each inter-turn sequence can also be seen in the fifth column in Table 8.1: Four of the linear units feature encapsulations and four feature overlays. As established in section 4.3.3.7, encapsulation can be divided into three categories according to the degree of explicitness: explicit encapsulation; fuzzy encapsulation; and implicit encapsulation. With reference to Table 8.1, of the four cases of encapsulation one is a fuzzy encapsulation: *this was the 70s* (line 13). As a fuzzy encapsulation, the deictic referent is the pro-form *this.* Such an encapsulation, as Moreno (2004) points out, leaves the reader with a fair degree of work to do to try to establish the exact stretch of the previous turn being encapsulated as the pro-form provides only an inexact reference. In this case, the

reader will therefore relate it to something relevant to the 1970s in cinema in the previous turn, such as the relatively lower technology. Thus line 6 *the effects are extreemly low quality* would seem to be a point of connection. However, as can be appreciated, fuzzy encapsulations very often remain somewhat vague as to what exactly they encapsulate.

The other three examples of encapsulation in the extract are implicit encapsulations (Lines 26, 28 and 34). In the case of the fuzzy encapsulation in line 13, it proved difficult, albeit eventually possible, to decide on a linear unit which could best be described as being encapsulated. In the cases of implicit encapsulation in the extract, it is not only difficult to find a linear unit which would be best described as being encapsulated, it is possibly erroneous to do so. Each of the examples would seem to provide a general evaluation of the whole of a previous turn. When Poster C comments *Wow, a kid "discussing" movies*, it would seem to be an evaluation of the whole of Turn 1. In terms of the linear analysis in this study, this means that such implicit encapsulations are described as simply following the default linear sequence, i.e. following on from the last linear unit of the turn in question, carrying, as it does, the role of the text of the moment. Such implicit encapsulations, then, provide a vaguer, more global evaluation of the previous discourse than may be the case with an explicit encapsulation.

Turning to the examples of overlay in the extract, we can see that there is a degree of variation in terms of how much of the previous linear unit is modified and to what extent the overlay is explicitly linked to the previous discourse. Turn 3 in line 24, for instance, contains an explicit reference to the point in the previous turn through the use of an

explicit signal *when i said*, and the near *verbatim* repetition of the previous linear unit *i didn't scare me but i enjoyed it*. As it is the same poster following up his/her own turn, it would seem to be a clear example of overlay being used as a means of clarification when it was felt by the poster that communication had not been as successful as expected.

The other examples of overlay in the extract are not explicitly signalled and are best described as 'acts of selective reference' (Sinclair 1993/2004:95), i.e. containing a reference to only part of the previous linear unit. In Turn 10 in line 39, the poster modifies the original phrase in line 1: *the most overated sci fi of all time* to make the more general comment *"the worst of all time"*. This modification is in itself part of the criticism of the original statement, in that Poster E is implying that the type of negative statement seen in line 1 is not original as it is a sentiment which has been expressed frequently. In the overlay in line 30, Poster E adapts the phrase *one of those AVP fanboys* in line 29 to *one of those trollboys*. This is a clear rephrasing of both the noun phrase structure (*one* + *of* + *those* + plural noun) and the morphology of the compound noun *fanboy* to coin his/her own compound noun *trollboy*. In the overlay in line 32, Poster A modifies the statement made in line 30 *More like one of those trollboys*, to create the question *How am I a troll?*

8.2.5 Discourse reflexivity in the *Alien* thread

The relative occurrence of discourse reflexivity may be seen as reflecting the shifts in focus of the topic of the thread. DR elements do not feature at all in the initial turn (Turn 1). In other words, there is no explicit reference to the present discourse in the initial turn.

This would seem to indicate that the poster's focus is exclusively on the expression of the negative evaluation of the film, i.e. as an *informer* on the autonomous plane rather than as *a text-constructor* on the interactive plane (Hunston 2000). Turn 2, similarly, is almost entirely focused on responding to the original turn in terms of the subject matter of the thread. There is, however, one occurrence of discourse reflexivity in line 15, an instance of the element **|oi (initial) *I MEAN*** as seen in section 6.5.3.3. In contrast to subsequent turns, however, this has a rather peripheral role in the turn. It does not occur at the beginning of the turn and so is not directly involved in the initial expression of disagreement. Instead, it provides an interactive signal in the local context of lines 15–16 that some sort of rephrasing or change of tact is about to occur (Mauranen 2012). This peripheral role of discourse reflexivity here reflects the fact that the turn is focused on engaging with the original post in terms of the film in question as opposed to the discourse itself.

Example 8.1b[111]

P	T	Line	Linear unit
A	1	1.	This is the most overated sci fi of all time!
		2.	I cant believe so many people liked this movie!
		3.	It is so boring!
		4.	The acting was horrible
		5.	(he's a f*#%ing robot!")
		6.	and the effects are extreemly low quality.
		7.	It was obviously a low budget film
		8.	and it deserves none of the credit it is getting .

[111] DR elements are indicated by being shaded and in bold. Linear units which contain DR elements are divided into the constituent elements, indicated with a forward slash between each element.

		9.	I rented this movie expecting a blockbuster
		10.	and I get a cheezy made for tv callibur snoozefest .
		11.	What is so good about this film?
		12.	This belongs with battlefeild earth and troll 2.
B	2	13.	um this was the 70s,
		14.	nothin super high-tech,
		15.	**i mean,** / yeah,/ its not like/ its using KingKong cgi...
		16.	i think u need to re-think this.
		17.	NOTHING belongs with troll 2. NOTHING.
		18.	it didt scare me,
		19.	but i enjoy it alot
		20.	if u hate alien, watch AvP,
		21.	invite ur friends over, pop some popcorn and make fun of the entire movie.
		22.	sorry tha u didnt enjoy alien,
		23.	me, myself, i prefer Aliens
B	3	24.	o and /**when i said**, / i didnt scare me / but / i enjoyed it, / **i ment Alien,** / not troll 2,
		25.	tha movie sucks
C	4	26.	Wow, / **a kid "discussing" movies.**
		27.	Incredible!
A	5	28.	Just because / I dont agree with you / does not mean / you have to / **resort to name calling**.
D	6	29.	He must be one of those AVP fanboys.
E	7	30.	More like / **one of those trollboys**.
		31.	**If you click on / his screen name**, / you'll find similar posts / by him/ on other boards.
A	8	32.	**How am I a troll**?
		33.	Oh I see, I have an opinion that you dont agree with and that I feel strongly about.
F	9	34.	**Shut the f uck up** / <name A.>,
		35.	you're out of your element.
		36.	I did not watch my buddies die FACE DOWN IN THE MUCK
		37.	this f ucking muppet, /this f ucking creep / **could critizize /** one of the best films ever made...

		38.	the movie whose structural perfection is matched only by its sheer awesomeness.
G	10	39.	If I had a dime / for every time / **someone called any movie/ "**the worst of all time!" / I'd be a millionare / by now.

The other occurrences of discourse reflexivity in the extract are all M linear units and are much more central to each turn. In Turn 3 in line 24, the linear unit contains a DR element **|m (initial)** ***WHEN I SAID***. This would seem to be similar in use to those |m (initial) element seen in section 6.5.2.3 such as, **|m (initial)** ***WHAT I'M TRYING TO SAY/IS*** **(+ot–** ***that***), where the role of the DR element is to try to maintain the self-consistency of what the poster is writing. The sequence **|m (initial)** ***WHEN I SAID*** **(+ot** ***that*****)** has the added quality of signalling to the reader that the immediate linear sequence may have to be flouted to find the exact relevance of the upcoming clarification (as signalled by the past tense *said*), i.e. that the point of connection is not line 23 but rather line 18. Here discourse reflexivity is used to guide the reader to the stretch of discourse in the previous turn *when i said* + quotation and then to signal to the reader with *i ment Alien* [sic], that a clarification is being provided as an 'act of consideration' (Mauranen 2010:24). In contrast, the other uses of discourse reflexivity in the thread reflect the ongoing dispute. For example, the discourse reflexivity in Turn 4 in line 26 and Turn 6 in line 29 express a somewhat ironic response to the original post and the original poster's behaviour. The first of these conveys a negative evaluation of Turn 1 by describing Poster A using the derogatory term *kid*. He also uses scare quotes with the DR verb *"discussing"* to describe Poster A's actions in the initial turn. The use of scare quotes

here, indicate the use of irony, i.e. that the intended meaning is not that which is normally used indicating a negative evaluation of Turn 1.

In Turn 9, discourse reflexivity is used in a seemingly insulting way in that it is used in to tell the poster to *shut the f uck up*. However, this may be somewhat dissipated by the fact that the poster is in fact adapting a famous speech by a character in a Coen brothers' film. Thus, it could be argued that discourse reflexivity is used here as part of the ironic humour expressed in the turn, which is prevalent in the IMDb corpus discourse. These issues will be discussed further in section 8.3.

In Turn 5, in line 28, discourse reflexivity is used as a means of protesting against Poster C's ironic criticism of the original post by referring to it as *name calling*. In doing so, Poster A criticizes Poster's C's behaviour as being inappropriate. This would seem to reiterate the finding in section 5.4.1 regarding the tendency of DR elements to be used to provide negative evaluation of discourse events. Turn 6 does not contain any DR elements and hence represents an attempt by Poster D to bring the discussion back to a criticism of Poster A for his cinematic tastes rather than his discourse behaviour. However, this is only a brief respite from the discussion about Poster A's behaviour as the thread moves to an accusation that Poster A is a *troll*. As we have seen in Chapter 5, *troll* can be considered as being a CMC-specific DR term, referring as it does to a poster who is guilty of committing personal attacks or behaving in a manner likely to provoke a negative response (Arendholz 2013:48). In the terminology of this study, then, a troll is a poster whose contributions are a deliberate attempt to provoke other posters to use a

suspensive response. In this case, Poster E rephrases Poster D's turn to coin the DR term *trollboy* to describe Poster A in line 30. This, in turn, is adapted by Poster A in Post 8 in line 32 to its more conventional form *troll* when s/he expresses his indignation in the question *How am I a troll?*

8.3 The role of discourse reflexivity in suspensive inter-turn M sequences

8.3.1 Introduction

In this section, we will further examine issues related to discourse reflexivity in suspensive inter-turn M sequences by re-examining the examples presented in the *Alien* thread and relating them to further examples from the IMDb corpus. Issues regarding the first linear unit of the sequence will be discussed in section 8.3.2, in particular those that occur in initial turns of the thread. In section 8.3.3, the second linear unit of the sequence will be discussed with particular focus on the co-occurrence of discourse reflexivity and the retrospective supplementary mechanisms of encapsulation and overlay in these suspensive responses.

8.3.2 Discourse reflexivity in initial turns

The initial turn in the thread inevitably plays a critical role in the development of the thread. It is in the initial thread that the subject to be discussed is proposed. It sets the tone for the thread and constrains the way that subsequent posters are likely to contribute. In terms of the linear analysis of this study, the initial thread provides a series of linear units from which subsequent posters can choose to be the first linear unit of an inter-turn sequence. The second poster then creates an inter-turn sequence by providing the second

linear unit of the sequence. If the initial poster includes an M– pro inter-turn linear unit, i.e. an elicitation at some stage, that in itself provides a clear point to which a subsequent poster can respond, thus creating a compliant M– /+ M sequence or suspensive M– / |M or |OI sequence. If there is no M– pro inter-turn linear unit, then the second poster is obliged to create a compliant M /MS or OIS sequence or suspensive M / |M or |OI sequence to respond to the previous turn.

This is, of course, true for any turn at any stage of the thread. However, initial turns have been chosen as the focus here, as a high number of such suspensive inter-turn sequences occur between the initial turn and the response to it. In this section, therefore, we will examine three initial turns. Firstly, we will re-visit the initial turn in the *Alien* thread described in Example 8.1a. This will be presented as an example of an initial turn which does not contain an elicitation and which provokes a series of rather antagonistic responses. Secondly, we will present the initial turn of another thread, Thread 013, the *Spacey* thread (see Example 8.2a), which contains an elicitation but one which seeks confirmation rather than information and which also provokes a number of antagonistic responses. A third initial turn from Thread 002, the *Audrey* thread (see Example 8.3), will be presented as an example of an initial turn which contains a genuine elicitation and which seemingly reflects a concern, on the part of the poster, to avoid subsequent suspensive sequences in the thread.

Firstly, then, we will examine the initial Turn (A1) of the *Alien* thread (see Example 8.1c). The most salient feature of this turn, as established in section 8.2.2, is that the initial turn

does not contain an M– pro inter-turn linear unit, i.e. it is an initial turn which does not contain a genuine elicitation. As was also established above, neither does it contain any DR elements, thus reflecting the prominence of the autonomous plane (Sinclair 1981/2004c) and of the poster in his/her role as informer at this stage for the discourse rather than text-constructor. This is confirmed if we consider the linear structure of the turn, containing, as it does, twelve M linear units. It is, therefore, a turn in which overt interaction through OI linear units is absent. Indeed, it is a turn in which oi elements within the linear units are similarly absent. Hence, the organizational elements present in the thread are exclusively textual in nature, thus adding to the sense that this is a poster who is concerned with the expression of opinion rather than in overtly interacting with other posters.

The subject line in line 1 is a core M linear unit and therefore a statement which is complete and which does not prospect completion from other posters. In non-technical terms, the poster is therefore stating an opinion rather than inviting contributions from prospective participants. It is not unusual for a subject line to be a core M linear unit as opposed to a pro inter-turn M– linear unit, with 28 of the 71 threads in the corpus beginning with a core M linear unit as the subject line (see Appendix 4.1).

Example 8.1c

T	L	Linear unit	Linear unit type
1	1.	This is the most overated sci fi of all time!	M
	2.	I cant believe so many people liked this movie!	MS– (Overlay)
	3.	It is so boring!	+M

	4.	The acting was horrible	MS– (Implicit encap.)
	5.	(he's a f*#%ing robot!")	|M (Overlay)
	6.	and the effects are extreemly low quality.	+M
	7.	It was obviously a low budget film	M–
	8.	and it deserves none of the credit it is getting .	+M
	9.	I rented this movie expecting a blockbuster	M–
	10.	and I get a cheezy made for tv callibur snoozefest .	+M
	11.	What is so good about this film?	M
	12.	This belongs with battlefeild earth and troll 2.	MS (Overlay)

It would be an overstatement to argue that because a core M linear unit is used in the subject line, an antagonistic tone is set for the rest of the thread. After all, a core M linear unit can just as easily be complimentary about the subject matter as critical. However, in the case of line 1 in Example 8.1c, a negative evaluation of the film to which the message board is dedicated combined with the fact that it is a linear unit which does not invite a completion would seem to at least augment the potential for antagonism in subsequent turns.

The rest of the turn is characterized as being a series of intra-turn linear units including three overlays (lines 2, 5 and 12), each, in general terms, reformulating the negative evaluation expressed in the previous linear unit. Although, as stated above, it is quite common for the subject line of a thread not to include a pro inter-turn M– linear unit, it is more unusual for the initial turn not to include any sort of elicitation at some stage. Around two-thirds of threads in the corpus (47 of the 71) contain at least one M– pro inter-turn linear unit.

The initial turn in Example 8.1c does contain a linear unit: *What is so good about this film?* in line 11 which, at first sight, appears to be an M– pro inter-turn linear unit. If it were an M– pro inter-turn linear unit it would certainly be an example of a *marked proposal* (Francis and Hunston 1992), i.e. an M– pro inter-turn linear unit which prospects not only a response but also the polarity of the response in order for the sequence to be completed. In this case the compliant response would have to be something like:

A: *What is so good about this film?*

B: *Nothing*[112]

in order to complete the discursive expectations created by the initial linear unit of the sequence. Such a response, however, would seem to be superfluous to the point that its inclusion may be interpreted as being somewhat ironic. It would therefore seem more accurate to classify line 11 as a core M linear unit which, by definition, does not contain an overt linguistic means to relate it to the previous linear unit nor does it prospect a completion. Rather than inviting completion, then, the linear unit in line 11 represents another statement of the initial negative evaluation at the beginning of the turn. In this way, line 12 can be interpreted as an overlay, again reiterating this negative evaluation by categorizing the film as belonging with two other science fiction films which are generally regarded as being two of the worst films of the genre.

[112] Invented example

As established in section 6.5.2.3, the use of such a stand-alone coda, i.e. a core M linear unit (or in this case an M / MS sequence) is a frequent way by which an extended turn is completed in the IMDb corpus discourse by acting as both a reiteration and a statement apart from the rest of the turn. However, despite the absence of an overt signal it can be argued that these codas play a role on the interactive plane by signalling to the reader that it is the end of the turn. Given the often evaluative nature of these final core M linear units, this can be related to previous claims that evaluation often occurs at boundary points in discourse. As Thompson and Hunston (2000:11) argue, 'evaluation at the end of each unit…marks that a point has been made and the reader's acceptance of that point is assumed.'

Thus, the combination of an unengaged M linear unit in the subject line expressing a negative evaluation, the prevalence of evaluative overlays reiterating the stance and the presence of a core M linear unit acting as a coda implying the acceptance of the evaluation of the reader can all be said to make it more likely that the response of subsequent participants is to be negative and possibly antagonistic and therefore that the responses would contain turn-initial suspensive linear units. In terms of the linear analysis of this study, then, all of this increases the likelihood of the linear units in the initial turn subsequently forming an inter turn M / |M sequence with the response of another poster.

Example 8.2a[113]

[113] This example is taken from Thread 013 about American actor Kevin Spacey, hereafter known as the *Spacey* thread.

P	L	Linear unit	Linear unit type (Supp. Retro Mechanism)	Element Structure of DR linear units
A	1.	hes not gay / is he?	M–	
	2.	i like him alot /as an actor	M–	
	3.	but / recently / my boy told me / hes gay / and / has never been married...	+M	
	4.	**someone please tell me** / its not true	**MS–** (Fuzzy encap.)	**|m/** m
	5.	not that / theres anything wrong / with it	|M (Implicit encap.)	

This can be compared to a similarly provocative initial turn of another thread (see Example 8.2a). In contrast to the *Alien* thread, the initial turn in the *Spacey* thread does contain a question. It is, however, overtly provocative regarding the supposed homosexuality of the actor. It should be noted that the alleged homosexuality of famous personalities is one of the most frequent topics in the IMDb corpus data. This thread is one of several in a similar vein about actors, such as Cary Grant, Marlon Brando, James Dean, Tom Cruise, Adam Sandler and Sacha Baron Cohen. The word *gay* occurs in 66 elements in the IMDb corpus and the various terms related to *homo** occur in 38 elements. This gives a combined frequency total of 104. Given that the word *actor** has a similar frequency of occurrences, featuring in 129 elements in the IMDb discourse corpus, it can be seen how frequently this topic arises. These threads are almost inevitably antagonistic and often draw accusations of inappropriate behaviour.

In this thread, then, the initial post begins in line 1 with a question tag or, in the terms of this study: a pro inter-turn M– linear unit constituting a core m element *hes not gay*

followed by an ois element *is he?* As Biber et al (1999:208) point out, such question tags generally elicit confirmation or agreement rather than elicit information. It is therefore an elicitation, albeit one that seeks clarification as opposed to information. It is, in Francis and Hunston's terms (1992:140), a marked proposal, in this case, prospecting a confirmation. It is therefore considered to be a pro inter-turn M– linear unit. An expected completion of the prospection would be something like:

A: *hes not gay / is he?* M–

B: *Yes, I'm afraid he is.* [114] +M

As Biber et al, also point out, such question tags are very often used as a means of echoing or of drawing a conclusion from a previous statement or exchange in conversation and are therefore the end result of a developing argument. In this case the logical order is reversed, the question tag comes before the explanation in line 3. This creates an abruptness in the turn adding to the overall sense that the stance provided in the subject line is non-negotiable and that the only acceptable way to respond to the question tag is by confirming its veracity. The fact that confirmation only is being sought by the poster is provided in line 4 with the rather arch request: *someone please tell me its not true.* This comprises a |m / m structure where the initial suspensive element *someone please tell me* is discourse reflexive. It is the point where the poster is explicitly requesting participation from other posters and therefore represents a linear unit where the interactive plane is prominent. However, it should be noted that it is a rather overdramatic means of rephrasing the initial elicitation in line 1. It therefore has the

[114] Invented example

effect of reiterating that a confirmation is the only expected response and thus adding to the provocative effect of the turn rather than opening the thread up to genuine discussion. Despite the similarities in the provocative nature of the initial turns in the *Alien* and *Spacey* threads, structurally, the initial thread in the *Spacey* thread is distinct from the initial turn in the *Alien* thread in that it contains a pro inter-turn M– linear unit. This means that any response to the turn is likely to respond to this linear unit forming either an M– / +M sequence if the response is compliant or an M– / |M sequence in the case of a suspensive sequence. In the event, eleven posters respond to the initial turn, ten of which begin with a suspensive retro inter-turn linear unit forming M– / |M sequences in the process.

The two initial turns seen so far can be contrasted with an initial turn that contains a genuine elicitation and a heavy use of DR elements as seen in Example 8.3. This is the initial turn of a thread taken from a message board dedicated to Marilyn Monroe in which the open question regarding which of Audrey Hepburn or Marilyn Monroe other posters think was more successful is posed.[115] The inclusion of an elicitation in the initial thread does not necessarily mean that the thread will be less antagonistic as we have seen in Example 8.2a. Nevertheless, in the case of the initial turn in Example 8.3, it can be seen that the assumptions made regarding the reader's acceptance of a stance as we saw in the initial turns in the *Alien* and *Spacey* threads are not evident here. It is therefore less likely to provoke an antagonistic response than the other initial turns seen above and therefore more likely to lead to compliant inter-turn M– / +M sequences.

[115] This example is taken from Thread 002 about American actor Audrey Hepburn hereafter known as the *Audrey* thread.

Example 8.3

P	L	Linear unit	Linear unit Type (Retro Supp. Mech.)	Element Structure of linear units containing DR elements
A	1.	Audrey or Marilyn?	M–	
	2.	Who was more successful?	|M– (Implicit Encap.)	
	3.	This isn't supposed / **to start a debate,**	**|M–** (Fuzzy Encap.)	m– / **+m**
	4.	i just want your opinions,	+M	
	5.	**you can explain** / if you want,	**M–**	**m–** / +m
	6.	but / **please dont diss / other posters**.	**+M**	**+ot– / m– /+m**
	7.	**I only put this topic**/ because / I was talking about Audrey / in my Visual Arts class	**M–**	**|m** / +ot–/ m /ms
	8.	and this girl didn't know who she was.	+M	
	9.	Then my friend said she was almost as big as Marilyn Monroe.	MS	
	10.	I however, thought Audrey was "bigger" or more successful.	MS (Implicit Encap.)	
	11.	I think they are both amazing.	M–	
	12.	One a sex symbol and one a fashionicon,	+M	
	13.	which one do you prefer?	MS– (Implicit Encap.)	

The way in which the poster in the *Audrey* thread seems to conceive his/her role in the thread may also be seen to affect the responses. S/he would seem to see him/herself as some sort of chairperson for the discussion about to begin in the thread. This type of role has already been noted in previous research into online message boards as the poster acting as the 'host' (Marcoccia 2004:135). This is in stark contrast to the initial poster in the *Alien* thread who treats the message board as a means of publically stating his/her stance. The initial poster in the *Audrey* thread's perception of him/herself as a chairperson is reflected particularly in the use of discourse reflexivity in Lines 3–7, something which

we have noted was absent from the initial in the *Alien* thread. Lines 3–6 constitute an explicit statement of the acceptable and unacceptable behaviour in the thread. In terms of relations between linear units, the linear unit in line 3 *This isn't supposed to start a debate,* is a suspensive intra-turn M– linear unit. Thus, the poster chooses to suspend the inter-turn prospection created in line 2 in order to explicitly state what is expected in terms of the responses. It is an M– linear unit as it prospects some sort of elaboration of what *this* actually is if it is indeed not a *debate*. This is a clear example of a series of intra-turn linear units in which the interactive plane is prominent with the poster in the *Audrey* thread acting as a text-constructor, or at least negotiating the co-construction of the text with the other posters.

This would seem to be similar to what in previous linear descriptions of spoken discourse (Sinclair and Coulthard 1975; Francis and Hunston 1992) is referred to as a *metastatement*, i.e. a statement which attempts to structure the discourse prospectively in some way. In the context of Sinclair and Coulthard's original research this was conceived as being something expressed by a teacher in the initial phases of a lesson. There are certainly similarities between the role of the teacher and what the poster here seems to be attempting to do here. The use of linear units containing DR elements in this way would also seem to be consistent with its use by the chairperson of academic seminars (Mauranen 2012), i.e. as an attempt to organize 'the entire speech event'. In this light, Lines 3–6 represent an attempt by the poster to restrict the behaviour of subsequent posters through the use of linear units containing DR elements 'with consequences for other speakers' freedom of movement' (Mauranen 2012:175). Such an attempt does not,

of course, guarantee compliant behaviour by subsequent posters. However, what we can see in Example 8.3 is discourse reflexivity being used as an attempt to overtly control subsequent behaviour whereas in the *Alien* thread the initial poster displays no overt concern as regards how the other posters behave in the thread. In terms of the linear analysis of this study, the initial poster of the *Audrey* thread is attempting to control the subsequent turns in order to maximize the possibility of compliant M– / +M sequences.

Therefore, we have seen three distinct initial turns. The initial turn in the *Alien* thread does not contain a pro inter-turn M– linear unit and therefore sets up either M / MS sequences or M / |M sequences. It has been argued here that the latter is more likely due to its structure and subject matter. Similarly, the initial turn of the *Spacey* thread is also likely to set up suspensive sequences given the topic and the overall tone of the turn, albeit M– / |M sequences in this case due to the presence of a pro inter-turn M– linear unit in the initial turn. The final example from the *Audrey* thread would seem to be a careful initial turn trying to avoid setting up suspensive sequences through the explicit use of discourse reflexivity on the interactive plane.

In the next section, the three distinct types of |M responses in which discourse reflexivity plays a prominent role will be examined.

8.3.3 Discourse reflexivity in suspensive responses

8.3.3.1 Discourse reflexivity in impersonal antagonistic responses

In this section we will examine the role of discourse reflexivity in what will be termed *impersonal antagonistic responses*. These are suspensive retro inter-turn M linear units which have the following characteristics:

- They contradict or challenge the previous linear unit;
- They contradict or challenge in a manner likely to provoke a further suspensive response;
- They are addressed to posters other than the poster of the linear unit being referred to.

In order to do this, how discourse reflexivity is used in linear units which contain an encapsulation will be examined as well as how it is used in those linear units that contain an overlay. Firstly, an example will be presented in which encapsulation features in an impersonal antagonistic response and how discourse reflexivity is used as a key means by which the previous stretch of discourse is negatively evaluated in the current linear unit.

Example 8.1d

P	**T**	**L**	**Linear unit**	**Linear unit Type**	**Element Structure of linear units with DR elements**
A	1	1.	This is the most overated sci fi of all time!	M	
		2.	I cant believe so many people liked this movie!	MS– (Overlay)	
		3.	It is so boring!	+M	
		4.	The acting was horrible	MS– (Implicit encap.)	
		5.	(he's a f*#%ing robot!")	IM (Overlay)	
		6.	and the effects are extreemly low quality.	+M	
		7.	It was obviously a low budget film	M–	

		8.	and it deserves none of the credit it is getting .	+M	
		9.	I rented this movie expecting a blockbuster	M–	
		10	and I get a cheezy made for tv callibur snoozefest .	+M	
		11.	What is so good about this film?	M	
		12	This belongs with battlefeild earth and troll 2.	MS (Overlay)	
C	4	27	Wow, / **a kid "discussing" movies**.	**|M** (Implicit encap.)	**|oi / m**
		28	Incredible!	MS (Implicit encap.)	
G	10	39	If I had a dime / for every time / someone called any movie/ "the worst of all time!" / I'd be a millionare / by now.	|M (Overlay)	|m– / +m– / +m– +m/ m / ms

Turn 4 of the *Alien* thread, as seen in line 27 in Example 8.1d, is an example of an antagonistic impersonal response with a retro inter-turn suspensive M linear unit containing both an encapsulation and a DR element. It begins with the **|oi (initial)** ***WOW***. As described in section 4.3.2.7, these 'Janus-faced' |oi elements in initial position both encapsulate the previous discourse and preface the upcoming linear unit.

As can be seen in Table 8.2, of the 303 suspensive retro inter-turn M linear units in the IMDb corpus discourse, 120 are initiated by a |oi element. This represents almost 40% of all suspensive retro inter-turn M linear units, meaning that prefacing an inter-turn suspension with a |oi element is a common, marked form. There are slightly more suspensive retro inter-turn M linear units that contain an encapsulation with an initial |oi element, at 66 occurrences as compared to those that contain an overlay at 54 occurrences.

Table 8.2 Initial element category in suspensive retro inter-turn M linear units

	m	**% susp. M retro inter**	**|oi**	**% susp. M retro inter**	**+ot–**	**% susp. M retro inter**	***Total***	**% susp. M retro inter**
Encap.	64	21.12	66	21.78	9	2.97	139	45.87
Overlay	104	34.32	54	17.82	6	1.98	164	54.13
Total	**168**	**55.45**	***120***	**39.60**	***15***	**4.95**	***303***	**100**

There are 57 different items which feature as initial |oi elements in suspensive retro inter-turn M linear units. Table 8.3 shows the ten most frequent items. Although it would be overstating the case to say that these |oi (initial) elements signal that the upcoming M linear unit is suspensive, it is true that a small number of |oi (initial) elements, such as *For Christ's sake* and *what the beep,* would seem to unequivocally signal some sort of contradiction or negative reaction. Others are less clear if taken in isolation.

Table 8.3: |oi (initial) elements in suspensive retro inter-turn M linear units

	Encap.	**Overlay**	**Total**
well / Well, well, well	10	7	***17***
<Name>	4	4	***8***
I think	5	2	***7***
(oh) yeah / yes	4	3	***7***
Hey	2	5	***7***
no no / no /uh no / nah	1	4	***5***
Wow	**4**	**1**	***5***
(oh) (my) God	3	1	***4***
LOL	3	1	***4***
maybe	2	2	***4***

With reference to Turn 4 in the *Alien* thread, a tentative generalization about what **|oi (initial) *WOW*** prefaces can be made. As can be seen in Table 8.3, it is one of the most frequent types of |oi (initial) element initiating suspensive M linear units with retro-turn

orientation in the IMDb corpus. Despite the *Collins Cobuild Advanced Learner's English Dictionary* (2003) definition of *wow* as being used 'when you are very impressed, surprised or pleased', five of the six cases of **|oi (initial) *WOW*** in the IMDb corpus are the initial element in a suspensive retro inter-turn |M linear unit. This would seem to indicate that **|oi (initial) *WOW*** in the IMDb corpus discourse is more likely to be used with irony rather than the dictionary meaning. Irony, here, is used to refer to verbal irony, i.e. a statement in which the meaning is different for the meaning ostensibly expressed (Abrams and Harpham 2009). In the case of **|oi (initial) *WOW***, the more frequent meaning in the IMDb corpus is actually the ironic meaning of expressing a negative evaluation of the previous discourse and prospecting some sort of contradiction or expression of disapproval. In turn 4 in the *Alien* thread, for example, **|oi (initial) *WOW*** in line 26 is the initial element in a suspensive retro inter-turn M linear unit containing an implicit encapsulation in the form of a core m element *a kid "discussing" movies*. As discussed in section 8.2.4, being an implicit encapsulation, it does not include either a referring pro-form or determiner to guide the reader to the point of encapsulation. Instead, it makes reference to a discourse act "*discussing" movies*. Such an implicit encapsulation leaves the reader with a fair amount of interpreting to do to ascertain what exactly is being encapsulated. As was already argued in section 8.2.4, in the absence of a pro-form or determiner + noun phrase, the reader will fall back on the 'default' linear interpretation (Sinclair 1993/2004e), i.e. that the current linear unit follows on from the last linear unit of the previous turn and that therefore the current linear unit encapsulates that. In electing to use such an implicit encapsulation, instead of limiting the reader to a specific point of connection in the sequence, the writer leaves the responsibility to the reader to find the

object of evaluation. It is therefore a deliberate gap in the coherence of the discourse which the writer leaves the reader to bridge. In this case, the bridging of the coherence gap is made more complex for the reader by the fact that the linear unit in question in line 26: *a kid "discussing" movies* is heavily ironic. Much of the ironic humour of this turn is then achieved when the reader succeeds in bridging the gap in coherence himself by making the conceptual connection between the implicit encapsulation and the relevant stretch of the discourse.

The DR verb *discuss* generally has positive connotations in the IMDb corpus discourse. For example, later in the same thread, Poster E continues to discuss Poster A's behaviour, as seen in Example 8.4. In the extract s/he contrasts the appropriate and *respectful* discourse act of *discussing,* in line 2, i.e. talking about a topic in a detailed and considered way through looking at different aspects relating to the subject at hand, with the inappropriate discourse act of *flaming*, which was introduced already in Chapter 5 as meaning displaying hostility through insulting and offensive language.

Example 8.4

P	**Line No**	**Linear unit**	**Type**	**Retrospective supp. mech.**	**Element Structure of DR linear units**
E	1	**Once again,/ if you show up / on a movie messageboard /** and **/ flame a movie,/ expect to be flamed / in return.**	|M		**|oi / |m / ms/** +ot– / **ms / m / ms**
	2	**Discuss it / in a respectful manner**, /	MS–	Overlay	**m / ms**
	3	and you will be treated the same.	+M		

In the case of Turn 4 in Example 8.1d, however, *discussing* is written within scare quotes indicating that the reader should not interpret the use of the verb as meaning the default meaning. Instead, combined with the already established ironic use of the **|oi (initial)** ***WOW*** the reader can presume that the verb is being used in an ironic negative evaluation of Turn 1. Thus, it can be interpreted as meaning that Poster C feels that Poster A has not succeeded in writing about the film in question in a detailed and considered way. Again, this would seem to be consistent with the findings in section 5.4.1, i.e. that DR elements referring to a discourse event tend to provide negative evaluation.

Poster C does not elaborate on this, so the reader is left to try to interpret why such a negative evaluation of Post 1 has been given. It may be that Poster C is basing the evaluation on the fact that, as we have seen, the initial turn was quite long and so at first sight appears to be a well-developed argument but in actual fact is a turn in which the same negative evaluation is reiterated through a heavy use of overlay. However, the negative evaluation may also be related to the fact that Poster A is evidently young and possibly immature. This can be seen in Poster C's use of the word *kid*. The DR use of the word *kid*, i.e. when it refers to a poster in the discourse, is quite unusual in the IMDb corpus. When it is used, it usually occurs as an ois element, i.e. as a supplementary vocative. Its use as an ois can generally be described as patronizing, for instance, *that aint bad, kid.* Here it acts within an m element. Nonetheless, the meaning is similarly patronizing. Combined with "*discussing*" in scare quotes, we can assume that the negative evaluation of Turn 1 is related to the fact that the poster indicates that s/he had only recently seen *Alien* despite it being a very well known film from the 1970s. The

irony is then reiterated through the use of a non-DR implicit encapsulation *Incredible!* in line 27. As Gibbs (2000) argues, an ironic statement conveys a pragmatic meaning by alluding to the failed expectations of the writer/speaker and the audience. In this case, the expectation of a rational, mature and well-developed discussion of the film *Alien* has apparently not been met according to the evaluation of Poster C. If encapsulation can be envisaged as a packaging of the previous stretch of discourse (Francis 1994:86), in an inter-turn context encapsulation in antagonistic responses of the type seen in Turn 4 can be considered to be *aggressive re-packaging* of the previous discourse.

However, if it is antagonistic, then, Turn 4 can be termed *impersonal antagonism*. Unlike Poster B, Poster C does not address Poster A directly. Instead, s/he is 'playing to the gallery' by referring to Poster A in the third person. Therefore the addressee of this message is not primarily Poster A, but rather the other readers of the thread. Poster A has thus become the subject of discussion and the overhearer, to use Goffman's term (Levinson 1983:72). Such a move away from a direct address to the poster does not necessarily mean it is a less personal attack. In this case, for example, it is possibly even more offensive as Poster A is left reading about him/herself being discussed by others. S/he is, of course, not excluded from the discussion and may him/herself respond to the turn, as s/he does in line 32.

A similar use of impersonal antagonism in a suspensive inter-turn M linear unit containing an encapsulation can be seen in one of the responses in the *Spacey* thread seen in Example 8.2b. However, in this case the effect is even more abrupt structurally as it is

the second linear unit of an M– / |M sequence. In contrast to the M/ |M sequence above, the M– / |M sequence actually suspends a prospection. In terms of linear structure, then, the M– / |M sequence would seem to be a clearer challenge to the linear expectations constituting, as it does, a refusal on the part of the second poster to even comply with a pro inter-turn M linear unit's expectation that a response will be provided in order to reach the target state. In non-technical terms, this may be described as a refusal to answer a question.

Example 8.2b

P	T	L	Linear Unit	Linear Unit Type (Supp. Retro Mechanism)	Element Structure of DR linear units
A	1	1	hes not gay / is he?	M–	
		2	i like him alot /as an actor	M–	
		3	but / recently / my boy told me / hes gay / and / has never been married...	+M	
		4	**someone please tell me** / its not true	**MS–** (Fuzzy encap.)	**|m/** m
		5	not that theres anything wrong / with it	|M (Implicit encap.)	
B	2	6	Could be /that /I am just suspicious /by nature, /I am beginning to believe /that /**we**, /once again,/ **have a Troll investation**.	**|M** (Implicit encap.)	|oi/ +ot– / |m/ ms |m / +ot–/ **m–** / |m/ **+m**

Turn 2 in Example 8.2b is similar to Turn 4 in the *Alien* thread in that it uses a DR noun phrase in an implicit encapsulation to negatively evaluate the initial turn of the thread. In this case, however, the appropriateness of the initial poster's behaviour is being challenged rather than the quality of his contribution as was the case in Turn 4 in the *Alien* thread. This is done by referring to the current situation in the message board as

suffering from a *troll investation* [sic]. We have already seen in the *Alien* thread, that the term *troll* is used as a means of categorizing a poster's behaviour as inappropriate. In this case in Turn 2, the negative evaluation of the DR term *troll* is reiterated by the fact that it acts a premodifying noun for the noun *investation* [sic], which is usually used to describe an invasion by insects or vermin. Just as we have seen in Turn 4 in the *Alien* thread, Poster B in this example chooses to make this rather offensive accusation indirectly. The DR pronoun of we in *we, once again, have a Troll investation* presumably refers to the accepted discourse community of the Kevin Spacey message board, from which Poster B is attempting to exclude Poster A.

In order to see how Poster B positions him/herself in this suspensive linear unit it is also important to note that the encapsulation is preceded by an elaborate seven-element suspensive sequence: c*ould be /that /I am just suspicious /by nature, /I am beginning to believe /that* with the element structure: |oi/ +ot– / |m / ms / |m / +ot–. This is what Hyland (1998a, 1998b, 2000, 2005) describes as a hedging device or, more accurately, a series of hedging devices. Although not discourse reflexive in themselves, they are employed by the poster as a means of a counter weight to the harsh negative evaluation in the DR encapsulation to come (Mauranen 2010). The question remains as to the Poster B's motivation for this counterbalance. It may be in order to express caution in order to diffuse any conflict with Poster A. However, since Poster B has chosen to exclude Poster A, placing him/her in the role of overhearer, this would seem unlikely. If the addressees are the other posters on the message board then the motivation would seem to be related to Poster B establishing him/herself as rational and a credible member of the message

board community, i.e. not a *troll*. This, then, allows him/her to provide an antagonistic suspensive response and at the same time mitigates any subsequent accusations of similar inappropriateness on his/her part.

We can now turn to an example of an impersonal antagonistic response which contains both DR elements and an overlay. Whilst an antagonistic response with an encapsulation can be envisaged as being an *aggressive re-packaging* of the previous discourse, an antagonistic response with an overlay may be considered to be *aggressive rephrasing* of the previous discourse. In Turn 10 in line 39 in Example 8.1d, Poster G does not address Poster A directly but rather evaluates, in general terms, the type of statements made in the subject line in line 1. Just as we have seen in Turn 4 in Example 8.1d and Turn 2 in Example 8.2b, the audience of the turn is not specified but it is presumed to be the other posters reading the thread other than Poster A. It is only clear that this linear unit refers to anything in the previous discourse when we reach the fourth element in the *Alien* thread *"the worst of all time!"*, which is in quotation marks. This would normally indicate a verbal echo but here it indicates a paraphrase of the subject line in line 1 in more general terms, thus providing an overlay. It is only then that the reader realizes that the third element, the DR element *someone called a movie* actually refers to the discourse act which occurred in line 1. Overlay is used here as a means of aggressively and almost surreptitiously modifying the original linear unit to something which can be more easily criticized. The fact that line 38 is expressed in hypothetical general terms adds to the negative evaluation of the subject line by implying that it is not only an invalid evaluation of the film but it is also a stance which is not original. There is also a sense of irony

related to the distance between the points of connection of the sequence here. Turn 10 is an attempt by Poster G to take the discourse back to the subject line, despite the intervening turns. Figure 8.1 provides a sense of the amount of activity that has gone on in the structure of the discourse between the first post and when Poster G intervenes. By implication, then, not only is Poster G challenging the appropriateness of the topic of discussion, s/he is also challenging the validity of all the subsequent discourse that follows in the thread.

8.3.3.2 Discourse reflexivity in personal antagonistic responses

In contrast to impersonal antagonistic responses, personal antagonistic responses are suspensive retro inter-turn M linear units which have the following characteristics:

- They contradict or challenge the previous linear unit;
- They contradict or challenge in a manner which is likely to provoke a further suspensive response;
- They are addressed to the poster of the linear unit being referred to.

There are two turns which may be considered to be personal antagonistic responses in the *Alien* thread: Turns 5 and 9. These can be seen in Examples 8.1e and 8.1f respectively.

Example 8.1e

P	T	L	Linear Unit	Linear Unit Type (Supp. Retro Mech.)	Element Structure of DR linear units
C	4	26	Wow, / **a kid "discussing" movies**.	IM (Implicit encap.)	I**oi / m**
		27	Incredible!	MS (Implicit encap.)	

A	5	28	Just because / I dont agree with you / does not mean / you have to / **resort to name calling**.	lM (Implicit encap.)	lot / m / +m– / +m– / +m– / **+m**

As we have seen in section 8.2.4.3, encapsulation can be used by the poster to aggressively re-package a previous stretch of discourse. Turn 5 in Example 8.1e represents a retro inter-turn lM linear unit containing an implicit encapsulation which describes the discourse act *resort to name calling*. This provides a negative evaluation Poster C's turn and line 26 in particular. Thus, *a kid "discussing" movies* is now reclassified as *name calling*. Through the use of a nominalization of the discourse act, Poster A is suggesting that *name calling* is an established behaviour-type already. Thus, in the case of Turn 5, through the very phrasing of the DR element *resort to name calling*, Poster A attempts to establish that this new evaluation of Turn 4 is a typical type of behaviour. There is also the implication in the use of the DR noun *name calling* that Poster C's own behaviour is childish as *name calling* is generally associated with children's behaviour. Thus, Poster A is reclassifying Poster C's response as childish and so is applying the same criticism of Poster C as Poster C himself had used in Turn 4, i.e. that s/he is immature. Poster A is, in effect, hoisting Poster C by his own petard. Thus, Poster A has employed discourse reflexivity to move the dispute away from the ironic criticism of the quality of the initial post as seen in Turn 4 to his/her own ironic criticism of Poster C's behaviour as a participant in the discourse.

Example 8.1f

P	T	L	Linear unit	Linear unit Type (Supp. Retro. Mech.)	Element Structure of linear units with DR elements
A	1	1	This is the most overated sci fi of all time!	M	
		2	I cant believe so many people liked this movie!	MS– (Overlay)	
		3	It is so boring!	+M	
		4	The acting was horrible	MS– (Implicit encap.)	
		5	(he's a f*#%ing robot!")	|M (Overlay)	
		6	and the effects are extreemly low quality.	+M	
		7	It was obviously a low budget film	M–	
		8	and it deserves none of the credit it is getting .	+M	
		9	I rented this movie expecting a blockbuster	M–	
		10	and I get a cheezy made for tv callibur snoozefest .	+M	
		11	What is so good about this film?	M	
		12	This belongs with battlefeild earth and troll 2.	MS (Overlay)	
F	9	34	**Shut the f uck up** / <name A.>,	|M (Implicit encap.)	**m** / ois
		35	you're out of your element.	M	
		36	I did not watch my buddies die FACE DOWN IN THE MUCK	M–	
		37	so / this f ucking muppet, /this f ucking creep / **could critizize /** one of the best films ever made...	+M	+ot– /m– / |m– / **+m–** / +m
		38	the movie whose structural perfection is matched only by its sheer awesomeness.	MS (Overlay)	

Another type of direct antagonistic response containing encapsulation can be seen in those responses which use an imperative verb. Despite the absence of the second person in such cases as *Shut the f uck up* line 34 in Example 8.1f, it is a clear instance of what Mauranen (2012:175) calls *other-oriented metadiscourse*. It is also what has been termed a directive move in previous models of analysis (Sinclair and Coulthard 1975; Francis and Hunston 1992; Tsui 1994). Most examples of this type are reasonably straight

forward. For example, in Turn B (Line 6) in Example 8.2c, it is clear what is being encapsulated and how it is being evaluated.

Example 8.2c

P	T	L	Linear unit	Linear unit type (Supp. Retro Mech.)	Element Structure of DR linear units
A	1	1	hes not gay / is he?	M–	
		2	i like him alot /as an actor	M–	
		3	but / recently / my boy told me / hes gay / and / has never been married...	+M	
		4	**someone please tell me** / its not true	MS– (Fuzzy encap.)	|m/ m
		5	not that theres anything wrong / with it	|M (Implicit encap.)	
B	2	6	For Christ's sake,/ read the threads/ **before asking irrelevant questions**, /please !	|M (Implicit encap.)	m / ms– /+m / ois

Just as we saw in Turn 4 in Example 8.1e, the response begins with the **|oi (initial)** ***FOR CHRIST'S SAKE***. As argued in section 8.3.3.1, it would seem clear that **|oi (initial)** ***FOR CHRIST'S SAKE*** prospects a suspensive linear unit. Here it expresses clear exasperation towards the initial post. What follows is an explicit condemnation of the appropriateness of the original marked proposal through the use of a |M linear unit containing the DR elements: *read the threads/* ***before asking irrelevant questions,*** */please !* This is achieved through an implicit encapsulation of the discourse act: *asking irrelevant questions.* It has been argued by Thompson and Hunston (2000) that evaluation of the present discourse is commonly on the relevant-irrelevant or important-unimportant parameter. It can be seen that such an evaluation is being explicitly stated in a linear unit

containing a DR element here. However, the structure of the noun phrase used is also important in terms of the degree of antagonism expressed in the response. As Hoey (2000:33) argues, if a writer uses a noun phrase containing a premodifier such as an adjective, 'it is more readily regarded by the writer and reader as given information or common ground.' Other such evaluations seen in suspensive retro inter-turn M linear units which contain DR elements in the IMDb corpus include: *a tough question*; *nice astroturfing*; *a condescending, snide comment; very very hard question*, *resort to childish insults*; *such false accusations*; and *a knee-jerk visceral rejecting reaction*. The irrelevance of the question in Example 8.2c here is therefore not up for negotiation.

Added to this is the sense of exasperation in that the DR noun *question* is actually in plural indicating that either this poster has previously asked irrelevant questions or that such irrelevant questions are a common occurrence in this message board. It is due to the irrelevance and frequency of such questions that Poster B has decided to create a suspensive M– / |M sequence and refuse to allow the discourse to reach its target state. All of this makes this a very clear, direct and obviously antagonistic response to the perceived irrelevance of the topic of the thread.

However, there are cases in the IMDb corpus discourse where suspensive responses seemingly go beyond simply provoking a further suspensive response to simply being offensive. Turn 9 in the *Alien* thread in Example 8.1f would seem to be one such case. Being an imperative, despite the absence of the second person, it is a clear example of *other-oriented metadiscourse*. However, it is difficult to categorize it in terms of the

retrospective supplementary mechanisms presented in this study: encapsulation; overlay; or verbal echo. Although it would not seem that *Shut the f uck up <A1's name>* fits comfortably into being either an encapsulation, overlay or verbal echo, the fact that it communicates a strong negative evaluation of the previous turn by asking the poster to stop talking without the use of a referring pro-form or determiner would seem to indicate that it is an example of an implicit encapsulation. As has been established certain cases of implicit encapsulation are as minimal as *Mmm* and *Yeah*, which implicitly evaluate the discourse as proceeding appropriately. A suspensive linear unit with an implicit encapsulation therefore implicitly evaluates the discourse as not proceeding appropriately. *Shut the f uck up <A's name>* certainly does that and so can be categorized as an implicit encapsulation.

At first sight it would seem to be an example of an extremely antagonistic, even offensive response. Again, it represents a turn initiated by a suspensive retro inter-turn M linear unit containing a DR element. If the initial turn in the *Alien* thread is a case of *flaming*, as Poster E contends, then, the motivation of such a post would be to provoke some sort of aggressive reaction. If this is the case, then, Turn F, at first sight at least, would seem to be an example of discourse reflexivity being used by a poster who has 'taken the bait' to respond in kind.

There are certainly cases in the IMDb corpus of discourse reflexivity being used in what can be termed offensive exchanges, for example in Example 8.5.

Example 8.5[116]

P	T	L	Linear unit	Type	Retrospective supplementary mechanism	Element Structure of DR linear units
A	1	1	Assassin!	M		
		2	Bruce Willis plays a great assassin.	M		
		3	His roles in The Jackal, The Whole Nine Yards, and Lucky Number Slevin were all great.	MS	(Implicit encap.)	
		4	I was wondering if he's had any other roles as assassins.	M–		
B	2	5	yeh doin ur mom last night	+M		
C	3	6	**You keep talking** / like a b*tch	|M–	(Implicit encap.)	m/ ms
		7	and im gonna slap you like a b*tch	+M		

In the discourse in Example 8.5, Poster C reacts to an obvious attempt by Poster B to respond to the original post in an extremely insulting way. In doing so, Poster C employs a suspensive M linear unit sequence to present a more elaborate and more offence alternative to the rather more prosaic *SHUT/UP* seen in Example 8.1f to express the order to desist from talking. The question remains, however, to what extent, we should consider these examples as cases of extremely belligerent behaviour or, alternatively, as a form of banter or of augmenting a sense of camaraderie among participants (Arendholz 2013). This type of behaviour would seem to be unacceptable to the discourse community and to the administrators of the IMDb website and is explicitly referred to as prohibited in the IMDb notice board in the *Terms and Conditions of Use*. As Arendholz (2013:48) comments, a 'sword of Damocles' continually dangles above the posters in message

[116] This example is from Thread 003 about American actor Bruce Willis.

boards as the administrator has the power to delete such inappropriate messages, as can be seen to have occurred on several occasions in the *De Niro* thread (see Appendix 4.3). However, it has also been found that such a negative attitude to such behaviour is not universally held among users of such message boards (Moor et al 2010). In fact it was found that users thought that such behaviour was often amusing and simply part of the type of interaction to be found on such sites.

It is true that a sense remains in threads such as Example 8.5, that the trade of insults is part of a set of almost ritualistic behaviour patterns than being truly belligerent. Within such an exchange it may even be argued that the focus is not on a genuine exchange of information or even insults but on the type of language used, i.e. on the very discourse reflexivity of the turns themselves. In this way, this type of exchange may be considered to be close in motivation, if not in execution, to *flyting* in 16th Century Scots poetry (Parsons 2011), where two rival poets exchanged insults through the medium of elaborated and sophisticated literary devices or to freestyle battles in rap music where two rappers trade improvised insults (Johnson 2008). In both of these mediums, what is being judged primarily is not the message but rather the use of language in terms of its originality, innovation or humour.

Returning to Turn 9 in Example 8.1f, a similar emphasis on such reflexive use of language is certainly evident. The cinema-literate reader realizes by line 36 that the linear unit containing a DR element in line 34 is not a simple insult or order to desist from communicating but rather a paraphrase from a speech made by Walter Sobachak, a

character in The *Big Lebowski*, a well-known Coen brothers' Hollywood film. The paraphrase continues through to line 37 as Poster F adapts the original speech to refer to the context of a response to Turn 1. Thus, the turn works on two 'levels': the first as a literal suspensive M linear unit expressing disapproval and asking Poster A to stop communicating and the second an intertextual play on words of a famous screenplay. The primary motivation of such apparently insulting language is therefore to express wit and humour rather than antagonism.

Thus, we have seen that discourse reflexivity plays a central role in antagonistic responses in a variety of complex ways. However, it would be a simplification to say that discourse reflexivity is only used in suspensive responses in order to provoke further suspensive responses. On the contrary, discourse reflexivity is also used in courteous suspensive responses as will be seen in section 8.3.3.3.

8.3.3.3 Discourse reflexivity in courteous suspensive responses

In this section, we will look at the role of discourse reflexivity in courteous suspensive responses both through the use of overlay and encapsulation.

An example of a courteous suspensive response using overlay and containing DR elements can be seen in Turn 3 in line 24 in the *Alien* thread in Example 8.1g. The linear unit in line 24 represents a suspensive linear unit which contains both an overlay and two DR elements. One of the most typical functions of a suspensive inter-turn overlay in the IMDb corpus discourse is providing some sort of clarification of a previous stretch of the

discourse. In this case, this is signalled by **|m (initial)** ***WHEN I SAID*** **(+ot** ***that*****)** prefacing a *verbatim* repetition of a segment of the previous turn in Lines 24–25. The DR **|m (initial)** ***WHEN I SAID*** **(+ot–** ***that*****)** can therefore be considered to form the following sequence:

|m (initial) ***WHEN I SAID*** **(+ot–** ***that*****)** + **overlay** or **verbal echo**

It therefore, provides the reader with an explicit signal as to the point of connection between Turns 2 and 3 as described in section 8.2.4. It is a signal to the reader that the strict linearity is to be flouted at this point, and that in order to make coherent sense of the evolving text, the reader must go beyond simply relating the 'text of the moment' to the linear unit that is immediately before. In this way, it would seem that one of the major functions of linear units which contain DR elements of this type is precisely as an indicator that strict linearity is not being adhered to in this instance and that the reader is expected to trammel up information that has already passed into the shared information records of the autonomous plane.

Hyland (2005:140) also argues that such material is used by the writer based on the norms of the discourse community within which the writer is acting and therefore that the use of such linear units containing DR elements are used by writers 'to present their arguments, control their rhetorical personality and engage their readers'. Hence, as well as providing a signal for the inter-turn coherence between Turns 2 and 3, the linear units in line 24 also contribute to the sense that it is an unedited, unprepared spontaneous response. The poster would seem to be deliberately simulating spontaneous spoken

language. This can be seen in the use of several other features: the use of |oi (initial) elements. e.g. *um*; *yeah*; *I mean*; using spelling normally associated with informal CMC, such as text messaging, e.g. *u*, *ur* (Tagg 2012:53) and lower case *i*. This is most noticeable in Turn 3 itself, where the poster responds to his/her own response with a clarification. Turn 3 is also initiated by the |oi element *o and*, which it has been demonstrated is used in texting (Tagg 2012:114) to preface something suddenly remembered or realized. In this case it prefaces a clarification in the linear unit in line 24. Although this may be simply related to personal style or to the fact that it is a poster who is more accustomed to using other types of CMC, it is also true that this simulation of unprepared spontaneous language, of which discourse reflexivity plays a part here, would seem also to be strategic. The poster is seemingly aware of the potential conflict created by creating an M/ |M sequence. S/he therefore allows for the possibility that the reader may think that it is not a well thought-out post and so the contradiction put forward may not be taken seriously. It is, then, effectively an overall strategy to hedge the contradiction in order to save face and avoid conflict.

Example 8.1g

P	T	L	Linear unit	Linear unit type (Supp. Retro Mech.)	Element Structure of linear units with DR elements
B	2	13	um this was the 70s,	|M– (Fuzzy encap.)	
		14	nothin super high-tech,	+M	
		15	**i mean,** / yeah,/ its not like/ its using KingKong cgi...	| MS (Overlay)	**|oi** / |oi / m– / +m
		16	i think u need to re-think this.	M–	
		17	NOTHING belongs with troll 2. NOTHING.	+M	
		18	it didt scare me,	M–	
		19	but i enjoy it alot	+M	

		20	if u hate alien, watch AvP,	M–	
		21	invite ur friends over, pop some popcorn and make fun of the entire movie.	+M	
		22	sorry tha u didnt enjoy alien,	M	
		23	me, myself, i prefer Aliens	MS (Overlay)	
B	3	24	o and /**when i said**, / i didnt scare me / but / i enjoyed it, / **i ment** Alien, / not troll 2,	|M (Overlay)	|**oi / |m** / m / |ot / |m / **m** / ms
		25	tha movie sucks	M	

In contrast to this, DR elements are also used in the IMDb corpus data in cases where the posters seem to be deliberately trying to establish a more formal style through explicit encapsulation. For example, the response seen in line 8 of Example 8.6 is a turn-initial |M linear unit with a DR element as the second element of an m / ms element sequence.

Example 8.6[117]

P	T	L	Linear unit	Linear unit Type (Supp. Retro Mech.)	Element Structure of DR linear units
A	1	1.	**3.** / The entire situation/ brought Ricky's father's / homosexual feelings/ to the surface.	MS (Implicit encap.)	**+ot–** / m– /+m– / +m– /+m
		2.	The gay neighbors and then, especially, what he thought he saw and his feelings toward his son's supposed sexuality brought his feelings to the surface and forced him to recognize them.	MS (Implicit encap.)	
		3.	They all just boiled over when he kissed Lester.	M	
		4.	The rejection is what made him kill Lester.	MS (Implicit encap.)	
		5.	4. / You can't disregard the irony / of a character /like that / being gay.	MS (Implicit encap.)	**+ot–** / m / ms / ms / ms
		6.	This film is all about American life	M–	

[117] This example is from Thread 062 about the film *American Beauty* hereafter the *American Beauty* thread, which featured already as Example 5.7 in Chapter 5.

		7.	and the human condition in general and irony is a huge part of life.	+M	
B	2	8.	I have to disagree / **with your point #3.**	|M– (Explicit Encap.)	m /**ms**
			It was NOT the rejection that made Col. Fitts kill Lester.	+M–	
			It was the fact that he had revealed his true nature in error,	+M	

It is an extract of a thread which is characterized as containing some quite extended turns. Turn 1, for example, is the end of a turn comprising 20 linear units. As established in section 5.4.5, it is a thread in which features characteristic of academic writing, such as the use of enumerating DR +ot– elements in Lines 1 and 5 (Biber et al 1999:880), as featured in previous models of metadiscourse as frame markers or sequencers (see Chapter 3) are present.

As stated above, Turn 2 would seem to be deliberately courteous. Like Poster B in Example 8.1h, Poster B in this example is clearly aware of the tension created by a suspensive linear unit which contradicts a point made by the previous poster. This can be seen in the use of an explicit encapsulation in line 8, within which there is the modalized m element *I have to disagree*, and a DR ms element *with your point #3*. Discourse reflexivity is employed here to defuse potential hostility in two ways. Firstly, the DR ms element gives the reader the point of connection (see section 8.2.4), i.e. to line 5. Such a discourse function of discourse reflexivity has featured in previous models of metadiscourse, for example as endophoric markers (Hyland 2005) or phoric (Ädel 2006) (see Chapter 3). A similar function of discourse reflexivity has already been seen in Turn 3 in the *Alien* thread above, where the DR sequence **|m (initial) *WHEN I SAID* (+ot *that*)**

+ **overlay** or **verbal echo** provides the signal to the reader that the strict linearity is being flouted.

The second function of the DR element in this context can be related to more interpersonal motivations. The DR element here acts as a means of qualifying the expressed disagreement to one particular point, therefore implying that the other three points made by Poster A are sound. This gives the impression of someone engaging in a rational and precise discussion on the subject rather than a blanket disapproval. The precision created by the DR element is instrumental therefore in not only creating a more academic tone compared to the *Alien* thread but also in dissipating potential conflict.

In conclusion, the use of DR elements in Turns 2 and 3 in Example 8.1g as well as Poster B in Example 8.6, provide examples of what has been conventionally seen as its role in previous studies. However, they also represent almost opposing strategies: discourse reflexivity as simulating an unprepared response thus providing a self-effacing response in order to reduce the potential conflict caused by contradicting the previous turn and a more precise 'academic' tone in order to provide a more exact and limiting opposite opinion to the original poster.

8.4 Chapter summary

This chapter has illustrated the role of discourse reflexivity in its most prominent place in the linear structure of the IMDb corpus discourse, suspensive M linear units with retro

inter-turn orientation, in other words, in turn initial linear units which are linked back to the previous turn but which challenge the discourse expectations in some way.

It has been shown that when such linear units contain an encapsulation and a DR element the role of the linear unit is almost universally to provide negative evaluation. It has also been shown that explicit encapsulation provides a more precise, often more localized evaluation whereas implicit evaluation is vaguer and possibly more global. Without guidance from the writer to the precise referent, the reader is left with the responsibility to make the connection. It has been argued that such DR encapsulation in this context can be conceived as *aggressive re-packaging* of the previous poster's turn.

It has also been shown that such suspensive linear units which contain an overlay have two functions, the first as a provider of clarification. This was as expected as clarification in order to facilitate otherwise less than successful communication is established as a central function of discourse reflexivity (e.g. Ädel 2006, 2010). The second role of DR overlay is similar to that of encapsulation as it also provides a negative evaluation of a part or all of the previous poster's turn. It is argued that this can be conceived as being an *aggressive reformulation* of the previous poster's turn. This can take the form of a surreptitious rephrasing of the previous post.

All of this points towards discourse reflexivity playing a central role in some of the most salient features of the IMDb corpus discourse, namely the prevalence of antagonist responses, i.e. of negative evaluative statements about the previous post; the prevalence

of irony and humour (whether it be laughing with the previous poster or at him/her); the manoeuvring of other posters by the writer such that a previous poster becomes an addressee or an overhearer and the type of personal or impersonal antagonistic responses this creates; the setting of the role for the initial poster, whether s/he conceives himself as primarily an informer or a text constructor, i.e. using the thread to invite others to contribute or as a soap box to express his/her opinion. It can be argued that all of these aspects of the IMDb corpus have discourse reflexivity at their core.

CHAPTER 9

CONCLUSION

9.1 Summary of main findings

The first and most obvious achievement of this thesis is the presentation of a multi-faceted linear profile of the IMDb corpus discourse relating to both element and linear unit relations as well as the role of discourse reflexive elements in both of these.

In relation to the initial research questions posed in Chapter 1, we can conclude that in the IMDb corpus discourse, certain DR elements have a tendency to behave in different ways to their non-DR counterparts in terms of their relations between elements. We can also say that, in discourse, certain linear units which contain DR elements tend to behave differently in terms of their relations between linear units when compared to linear units that do not contain DR elements. To be specific the two main findings of this thesis can be summed up as follows. In the IMDb corpus discourse, it was found that:

- DR elements tend to be at their most prominent in suspensive elements, both message-oriented and interaction-oriented in the initial position of the linear unit.
- Linear units that contain DR elements are at their most prominent in suspensive linear units oriented towards the previous turn. This is particularly true for those that contain an encapsulation. These often, although not exclusively, occur at points where negative evaluation and/or antagonism is expressed.

The main objectives of the thesis, as noted in Chapter 1, were to provide a description of the IMDb discourse in terms of the linear relations between elements in general and DR elements in particular and in terms of the linear relations between linear units in general and those linear units that contain DR elements in particular.

The main finding from the element analysis is that the IMDb corpus discourse is predominantly message-oriented, as around three quarters of all elements are m elements with a large number of supporting ot elements and a number of oi elements. However, the most distinctive features of the IMDb discourse are seen in oi elements, most obviously CMC-specific elements and informal interactive devices, especially attitudinal and engagement markers. In contrast, m elements resemble other data from other corpora and ot elements are made up of a small number of highly frequent conjunctions. It was also found that a characteristic of the IMDb discourse was the large number of supplementary ms elements and suspensive message-oriented (|m) or interactional (|oi) elements.

As regards DR elements in element relations, over five percent of all elements in the IMDb corpus are DR elements, most of which are m elements. Although there is no significant difference between the number of DR m elements in general when compared to what would be expected if they were distributed in the same way as non-DR m elements in the corpus, there is a significantly higher number of initial suspensive |m elements (e.g. *I'd say, I'm just saying*) when compared to the number that would be expected if they were distributed in the same way as non-DR initial |m elements. This is also true for DR oi elements in general, especially |oi elements in initial position (e.g.

Let's say; *I mean, I guess*, *LOL*). It was also found that there was a significantly lower number of ot elements than would be expected if they were distributed in the same way as non-DR ot elements. It was shown that DR m elements which refer to discourse events and discourse acts very often show a negative evaluation.

Additionally, |m and |oi elements perform a variety of roles in the linear structure of the IMDb corpus, often depending on their turn orientation. Within a turn they often appear in linear units at the end of turns as a reiteration; or a means of signalling that the reader should go beyond the previous linear unit for the relevance of what is being written. In inter-turn orientation, they appear both in compliant linear units which signal conviviality and which support the previous poster's turn and, more often, in suspensive linear units in which they are better thought of as signals of antagonism and conflict.

The main findings as regards discourse organization in the IMDb corpus is that the vast majority of linear units in the IMDb corpus discourse are message-related (M) linear units. It was also found that the IMDb corpus is overwhelmingly retrospective rather than prospective in its orientation.

There are also a large number of linear units which are suspensive particularly in inter-turn orientation meaning that in over half of the times that a participant in the IMDb discourse responds to a previous turn, s/he challenges the validity of the previous participant's contribution. It was also found that in the IMDb corpus that the M / MS

sequence is around twice as frequent as the M– / +M sequence meaning encapsulation is around twice as frequent as prospection / completion. The most frequent retrospective supplementary mechanism present in M / MS sequences is encapsulation.

As regards the role of DR elements in the linear structure of discourse it was found that over 12% of linear units feature at least one DR element. Discourse reflexivity is particularly prominent in suspensive M linear units, especially in those that contradict or question the validity of the contribution of another participant. The results also show that discourse reflexivity is prominent in the second part of M / MS sequences and not in M– / +M sequences. In all of the categories in which discourse reflexivity is prominent, it is in encapsulation, especially implicit encapsulation, that it is the most prominent.

Furthermore, when such linear units contain an encapsulation and a DR element in the IMDb discourse, the role of the linear unit is almost universally to provide a negative evaluation. It has been argued that such DR encapsulation in this context can be conceived as *aggressive re-packaging* of the previous poster's turn whereas DR overlays in a similar context can be thought of as *aggressive reformulation* of the previous poster's turn.

All of this points towards discourse reflexivity playing a central role in some of the most salient features of the IMDb corpus discourse, namely the prevalence of antagonist responses, i.e. of negative evaluative statements about the previous post; the prevalence of irony and humour (whether it be laughing *with* the previous poster or *at* him/her); the

manoeuvring of other posters by the writer such that a previous poster becomes an addressee or an overhearer and the type of personal or impersonal antagonistic responses this creates and so on. In this respect, this study offers a myriad of possible related research areas, including:

- Whether discourse reflexivity has similar linear structural roles in other CMC interaction as those seen here in online message board discourse;
- Whether discourse reflexivity has a similar role in terms of establishing negative evaluation and possible antagonism in other types of dialogic and polylogic discourse.
- Whether discourse reflexivity has a similar role in academic CMC discourse (online message boards and beyond).

9.2 Implications of the thesis

9.2.1 The concept of discourse reflexivity

As well as providing specific insights into online message board discourse this thesis can also be conceived as contributing to the existing field of knowledge in more general terms. As described in chapter 2, discourse reflexivity and metadiscourse has attracted a huge amount of academic attention in the last twenty-five years or so. Like Mauranen (2010:14), my approach in this study has been to adopt a 'back to the basics' approach to the definition and treatment of the concept. As well as fitting in with the overall ethos of Sinclair (1992, 1993/2004e) and Sinclair and Mauranen (2006) to reduce the 'vast complexity of human communicative behaviour…to a small number of simple activities' (Sinclair 1992:83) seen in the linear model employed here, this approach also allowed the

research into the *structural* role of discourse reflexivity (as summarized in section 9.1) as opposed to the functional orientation normally associated with the well-established taxonomies of metadiscourse (e.g. Vande Kopple 1985; Hyland 2005; Ädel 2006) to be explored.

The concept of discourse reflexivity in this study is one which almost entirely focuses on the notion of the explicitness of the reference to the present discourse. This defining criterion had already been established, particularly by Mauranen (1993a, 2007, 2010, 2012) and Ädel (2006, 2010). However, due to the fact that much of the IMDb discourse, and discourse in general, is overtly interactive, then, any reference to the writer and the reader has to demonstrate their role as *text-constructor* and a *person informed of the construction of the text* respectively for the reference to be considered to be discourse reflexive.

Nevertheless, even with this more refined defining criterion, ambiguities remain. As was seen in section 6.5.2.2, the DR verb *say* conveys a literal meaning of communicating by speaking as well as more idiomatic meanings. It can, in fact, occur in all three main element category types:

Example 9.1: m element[118]

P	Line	Element	Element type
A	1	**How can you say**	**|m–**
	2	this is	m–
	3	the worst Oscar movie ever	+m

[118] This example is from Thread F10 about the film *Braveheart*.

Example 9.2: oi element[119]

P	Line	Element	Element type
A	1	Maybe	|oi
	2	**it is possible to say**	**|oi**
	3	that	+ot–
	4	Tom has never acted badly	m

Example 9.3 ot element[120]

P	Line	Element	Element type
A	1	This kind of	m–
	2	homoerotic behaviour	+m–
	3	should be included	+m–
	4	in the rating system	+m
	5	**Say**	**+ot–**
	6	"Language,	m
	7	violence,	ms
	8	raunchy same sex nudity."	ms

Example 9.1 is an example of a DR m element containing *say*. This example is clearly DR, referring back, as it does, to another poster's previous turn. The other two examples above are examples of DR o elements containing *say*. These are less clearly DR as they are less literal in terms of their reference to the present discourse. It can be argued that the DR |oi element *it is possible to say* refers to the upcoming speech act in line 4. However, it is also possible to argue that it represents a semi-fixed phrase which has gone through a process of pragmatic re-orientation (Butler 2008), whereby its pragmatic function, i.e. as a hedging device, now overshadows the original literal DR meaning. If the original DR

[119] This example is from Thread A38 about American actor Tom Hanks in which posters are asked what his worst performance is.
[120] This example is from Thread F01 about the film *Borat* and its supposed homoerotic content.

meaning has been semantically bleached, to paraphrase Hopper and Traugott (1993), to such an extent that it is not perceivable any more, then it cannot be considered an explicit reference to the discourse.

The process of pragmatic re-orientation would seem to be more advanced in the DR ot element *Say* in Example 9.3. In this example, the pragmatic meaning of providing an example would seem to have overshadowed the original meaning to the point that it is now difficult to extrapolate the literal meaning of *say* from the pragmatic use. The same could be said for other fixed expression DR o elements containing *say* such as the oi element *Let's say* and the ot element *needless to say*, in which the literal meaning has less importance and is more opaque than its interactional or pragmatic meaning. It would therefore seem fruitful to further explore the comparative roles of literal and idiomatic discourse reflexivity in the linear structure of grammar and discourse.

9.2.2 Linear Unit Grammar

The contribution of this thesis regarding Linear Unit Grammar (LUG) can be divided into its application and its elaboration. Since being published, as we have seen in Chapter 2, LUG has attracted some attention in terms of its application. However, the present study would seem to be the first corpus-based application of the LUG and the application of the model (albeit embellished) to a large amount of data and thus provides some sort of validation of LUG.

Additionally, this study has added to the original LUG model. This, it must be said, was done with caution. As has been stated previously, the general ethos of this study was to keep the models to a small number of activities. The first addition was the distinction between suspensive and compliant elements. This was partly based on the inclusion of suspensive elements in Brazil (1995) and partly based on the extension of the concept of *challenge* from discourse analysis. It thus provided a means of categorizing occasions where the linear expectations were not being met within the context of linear units. The second addition was the sub-categorization of o elements. The addition of suspensive elements allowed the system, for instance, to be able to classify those linear units which were initiated by an o element but which did not provide a retrospective link.

The present study has also proposed a means of advancing LUG towards the classification of elements and element sequences in a lexico-linear description. As was noted in Chapter 2, Sinclair and Mauranen encourage the combination of LUG with other established analytical frameworks. However, within the scope of elements itself, there is, it seems, much further research necessary to bring LUG forward to a point where aspects of phraseological descriptions, such as Sinclair's own Lexical Grammar (2003/2004g) can be incorporated into LUG in a systematic way. In section 6.5.2, this study proposes a means of combining the information gleaned from LUG categorization with the lexis involved to produce a lexico-linear characterization which most closely resembles a pattern (Hunston and Francis 2000) but which includes supplementary information now available from the linear description from this study. This lexico-linear characterization includes: the subcategory of the element in which the lexical item occurs; the lexical item

itself, including information regarding whether it is a fixed phrase or if potentially there are intervening words; and whether it typically combines with other words or elements.

In Chapter 6, several of these were presented including the DR **|m *I/SAY* (+ot– *that*)** sequence. With a corpus of only 40,000 words it was difficult to make any more conclusions other than that it was not a completely fixed sequence in that words could occur between *I* and *say*. In order to demonstrate the potential of this type of characterization we can look at the occurrence of **|m *I/SAY* (+ot– *that*)** sequence in a larger corpus compiled of IMDb online message board discourse of around one million words. From this characterization, we can see that the element is varied in the number of words between *I* and *say* and in the actual words that feature. The sequence featured 230 times in the corpus. As can be seen in Table 9.1, the most frequent distance between *I* and *SAY* in this sequence is I*SAY. The two most common distances, I*SAY and I**SAY, account for over 70% of all the sequences. These can be further divided between the literal DR meaning, i.e. 'tell through words' and the idiomatic meaning, i.e. an expression of a stance as discussed in Chapter 6. These are illustrated in Table 9.1.

Table 9.1 Distance and variation in |m *I/SAY* (+ot– *that*) sequence

Distance	**Examples**		**Total**
Meaning	**Literal**	**Idiomatic**	
I*SAY		*would; d; can; could; must; will; gotta; cannot; ll; ever; should; still*	105
I**SAY	*didn't; did not; don't; meant to; wanted to; were to;*	*have to; can't; can honestly; ll just; can definitely; can never;*	59

		can simply; can still; d also; d prolly; d rather; d really; get to; got to; shouldn't; still would; will also; will always; will definitely; will just; won't; would hardly; would never; would safely	
I***SAY	*am trying to; m trying to; did want to; didn't just; only wanted to;*	*would have to; can't honestly; d have to; d tend to; definitely have to; do want to; just have to; ll have to; m ashamed to; m prepared to; m proud to; m ready to; still can't; too have to; will never ever; would not really; wouldn't exactly*	31
I SAY		✓	28
I****SAY	*wasn't going to; don't mean to;*	*won't begin to*	3
I******SAY	*sure as hell not going to*	*would go as far as to*	2
I*****SAY		*d be hard pressed to*	1
I*******SAY	*personally don't have the audacity to*		1

Hence, in this way the model presented in this study may be further developed to incorporate such lexico-linear information in the characterization of elements and element sequences.

9.2.3 Linear description of discourse

This thesis has also contributed to the study of a linear description of discourse. One of the principal aims of this study was to take up the agenda laid down by Sinclair twenty

years ago for the description of discourse. This thesis has achieved this by operationalizing what were two rather explorative and sketchy frameworks (Sinclair 1992, 1993/2004e). Not only that, this has been done employing a corpus-based approach, which as far as this writer knows, is the first time that such a study has been embarked upon.

Perhaps the greatest innovation of this thesis is the integration of LUG and the outline of spoken discourse and the analytic system for written discourse into one unified system by employing the same categorization system of M and OI and their subcategories: M–, MS– etc. This gives future researchers the means of applying linear principles to both grammar and discourse to the same text. The inclusion of *suspension*, although not new to linear descriptions of discourse (as *challenge* in other related models), is rarer in the description of written discourse and in this study it has provided a unifying principle in the distinction between core, compliant and suspensive categories in elements and linear units.

The inclusion of turn orientation as a further means of classification also allows researchers to apply the model to monologic, dialogic and polylogic discourse. As has been seen in the summary in section 9.1, such an inclusion has been crucial in describing the exact nature of suspension in the IMDb discourse.

It is hoped that this study has allowed the reader to see the profound effect on the conception of structure that a dynamic view of discourse has. If the researcher puts

him/herself in the position of the reader focusing on the text of the moment, there are four possibilities. The text of the moment is retrospectively linked to the previous linear unit; it prospects a link to the following linear unit; it does both; or it does neither. A prospection (M–) prospects a completion (+M); if the previous linear unit does not prospect and the text of the moment creates a retrospective link with the previous linear unit then this link is in the form of an encapsulation, overlay or verbal echo (these are MS or OIS depending on whether the linear unit is message oriented or interaction-oriented). This means that there are two basic sequences in discourse:

M / (MS or OIS)

M– / +M

If we add the possibility that the second linear unit can be suspensive we then have the following four possibilities:

M / (MS or OIS)

M / |M or |OI

M– / +M

M– / |M or |OI

As we have seen in this study, other supplementary information can be added for greater delicacy in the description, such as the turn orientation and the type of retrospective supplementary mechanisms employed. This provides a powerful tool for the universal

application of linear descriptions in discourse for future researchers. The model proposed here for the study of discourse has thus both successfully reduced the 'vast complexity of human communicative behaviour…to a small number of simple activities' and allowed for the possibility of applying descriptive devices employing the same principles to produce results of the complexity seen in Chapter 7 or as detailed a qualitative description as seen in Chapter 8 of this study.

9.3 Limitations and future research

There are several limitations to the present study. First of all, as a purported corpus-based study, employing a corpus of 40,000 words inevitably means that any claims to generalizability are tentative. This is particularly true when we consider that the calculations of significance were carried out based on 15,000 elements and 4,000 linear units. A larger sample would have aided in this respect, although it must be said that this study is more extensive than most other studies employing *UAM CorpusTool* in comparable studies.

It would also seem to be the first time that the *UAM CorpusTool* has been employed for linear analysis. However, it should be noted that the *UAM CorpusTool* as it stands has several limitations in its use on a linear based description. It is, after all, a tool devised for the classification of paradigmatic descriptions of language, such as Systemic Functional Linguistics. It is therefore designed for the classification of several options in a system but not for the easy retrieval for information about syntagmatic sequences. A possibility therefore for future research would be to enhance such a tool as *UAM CorpusTool* in

order to be able to report, for instance, how many suspensive M linear units contain the sequence m / ms element or to provide all examples of an element sequence such as: ⌊oi elements / +ot– *but*.

Both Sinclair (1993/2004e) and Sinclair and Mauranen (2006) emphasize the individual nature of how each person processes and interprets the incoming language as a reader/listener. For that reason, Sinclair and Mauranen (2006) argue that inter-rater reliability studies of LUG would not be of great import. Nonetheless, as was outlined in section 4.3.2 above, a limited inter-rater reliability procedure was carried out in this study indicating that inter-rater reliability was, in fact, very high in the initial stages of the categorization of element relations, namely in the establishing of PUBs in the chunking of the text and the establishing of the broad distinction between m and oi elements. This provides some evidence of the reliability of the system of analysis employed in this study.

It should also be recognized, nevertheless, that a number of the decisions made both in establishing the PUBs and making the distinction between m and o elements in this study can be characterized as being marginal. This was particularly true of those initial ⌊oi elements that do not display the 'syntactic independence' (Müller 2005:5) characteristic of 'discourse markers' such as *well*, *I mean* or *you know*. In contrast, more syntactically integrated interactional elements such as *so what if* and *no matter* are not clearly separated from the rest of the linear unit. In an element such as *so what if* (see Example 9.4) it is unclear where the PUB should be established since both before and after the word *if* are seemingly equally plausible. Additionally, it could be categorized as an oi

element or an m element as it could be seen to be either an attitudinal marker signalling how the upcoming linear unit is to be taken (i.e. an oi element) or an integral part of the message (i.e. an m element). In the end, the form interpretation was applied in this case. However, it is recognized that this is a potential area of variability between raters in the system of analysis.

Example 9.4 [121]

P	L	Element	Element type
O	277	So what if	|oi
	278	his career	m–
	279	lately	|m
	280	hasn't been	+m–
	281	as great as before,	+m
	282	but	+ot–
	283	it happens	m

Further research into inter-rater variability in general and regarding the distinction between initial |oi and |m elements in particular would therefore be beneficial. This is particularly pertinent in this study as DR elements are at their most salient in initial |oi and |m elements.

Example 9.5[122]

P	L	Linear unit	Type
Q	315	De Niro is	m–
	316	the most admired actor	+m
	317	currently working today.	ms
	318	FACT.	oi

[121] This example is taken from Thread 017, the *De Niro* Thread.
[122] Ibid.

Neither is such subjectivity in the categorization of elements limited to initial |oi and |m elements. For instance, the distinction between core and qualified oi elements is also less than clear on occasion. In Example 9.5, the oi element *FACT* could be considered to be either a core oi element or a qualified ois element depending if it is considered to be separate from or supplementary to the previous sequence of elements. The decision to categorize the element as a core oi element as was take in this study means that it is subsequently categorized as being a separate OI linear unit as seen in Example 9.6 below rather than as part of the M linear unit in line 88.

Example 9.6[123]

P	L	Linear unit	Type
Q	88	De Niro is the most admired actor currently working today.	+M
	89	FACT.	OIS

Such an example as this suggests that the distinction between the two scopes of analysis, which has been maintained in this study, i.e. between relations within the linear unit and relations outside the linear unit (or between 'grammar' and 'discourse' respectively as they have conveniently termed) may not be as clear cut as has been implied thus far in the present system of analysis. Further research into whether the distinction between the two scopes of analysis can be consistently maintained and whether this is also an area for potential rater variability would therefore also be of benefit.

[123] Ibid.

The dynamic nature of the model also poses significant challenges for the researcher. Sinclair and Mauranen describe a process in which the reader/listener is continually hypothesizing what will come up. Such hypotheses are proved to be correct or disproven and forgotten as the ongoing processing of the discourse moves on. Intuitively, this would seem to be true. However, it places the researcher in a tenuous position, whereby his/her own prejudices, experiences and skills (limited or otherwise) are brought to bear on the text of the moment in order to decide if it is linked in any way to the previous unit or if it prospects completion. In the case of retrospective links, in discourse the analyst can pause and assess the presence of an encapsulation or an overlay in order to categorize reliability. However, the same cannot be said for prospection. Not only is there likely to be much more scope for different interpretations between researchers, but also the researcher may have two or more hypotheses as to how the text might continue. It may be true, as Sinclair (1992/2004a) argues, that prospection does not mean that the future text is determined but rather that the next unit is judged in light of the prospection. However, by then the prospection die has been cast for the researcher and, if a dynamic model is to be applied strictly, this decision should not be changed even in the light of the subsequent text. This has serious implications for the reliability of a dynamic linear description (as opposed to a synoptic one) in the potential variability and also raises the question of whether the discarded or disproven hypotheses should be equally recorded. In terms of the linear model here, then, the classification of prospection remains problematic. Fortuitously, the focus of this study was on discourse reflexivity which, as it turned out, is more salient in encapsulation than prospection. However, the reason that sub-categories for prospection were not proposed for the system of analysis as they were for

encapsulation in the system of analysis for this model was partly due to such problems of interpretation and hypothesizing. Further research into this area would seem to be pressing if a dynamic model of the description of discourse is to progress.

By the same token, the development of the category overlay would also be beneficial. As noted in Chapter 2, overlay is a complex mechanism which requires a more systematic analysis than the present study has permitted. Given that it was found in Chapter 8 that overlay is used in a wider range of functions and achieves more subtle effects than encapsulation, it would seem that more research specifically into the mechanics and the role of overlay would be beneficial.

A more comprehensive validation of the model in terms of actual reading practices would also seem to be necessary. This study has taken on face value that Sinclair and Mauranen's intuition as regards how we read and process information is accurate. However, the exact role of chunking in the processing of information is inevitably more complex than that stated by Aubrey. We do not, after all, read a text out loud in our heads in chunks. It would also be beneficial for more research to test Sinclair's hypothesis that in normal reading we do not go back and forth in the text. This assertion would seem to be most urgently tested on further text types in CMC other than online discussion boards, such as the type of webpage where hyperlinks and non-linear features of text are prominent.

However, despite these limitations, it is felt that the general objectives of this thesis have been achieved. This thesis has presented an innovative linear model for both grammar and discourse that promises to have a wide range of applications for future research into linear structure. Furthermore, we now have some base line figures which can be applied in comparative studies employing the same model in order to study similar phenomena in other contexts. Finally, we now have the means of re-appraising the concept of discourse reflexivity in terms of its role in linear structure.

REFERENCES

Aarts, J. and W. Meijs (eds) (1990) *Theory and Practice in Corpus Linguistics*. Amsterdam/Atlanta: Rodopi.

Abdollahzadeh, E. (2010) 'Poring over the findings: Interpersonal authorial engagement in applied linguistics papers' *Journal of Pragmatics* 43: 875–890.

Abney, S. (1991) 'Parsing by chunks' in Berwick R. et al (eds), pp. 257–78.

Abrams, M. H., & G.G. Harpham (2009) *A glossary of literary terms* (9th ed.). Boston: Wadsworth Cengage Learning.

Ädel, A. (2001) 'On the search for metadiscourse units' in Rayson et al (eds), pp. 3–12.

Ädel, A. (2005) 'On the boundaries between evaluation and metadiscourse' in Tognini-Bonelli, E. and G. Del Lungo Camiciotti (eds), pp. 153–162.

Ädel, A. (2006) *Metadiscourse in L1 and L2 English*. Amsterdam: John Benjamins Publishing Company.

Ädel, A. (2008) 'Metadiscourse across three varieties of English: American, British and advanced-learner English' in Connor, U., E. Nagelhout and W.V. Rozycki (eds.), pp. 45–62.

Ädel, A. (2010) 'Just to give you kind of a map of where we are going: A Taxonomy of Metadiscourse in Spoken and Written Academic English' *Nordic Journal of English Studies* 9/2: 69–97.

Ädel, A. and A. Mauranen (2010) 'Metadiscourse: Diverse and Divided Perspectives', *Nordic Journal of English Studies* 9/2: 1–11.

Aijmer, K. (2002) *English discourse particles*: *Evidence from a corpus*. Amsterdam/Philadelphia: John Benjamins.

Aijmer, K. and B. Altenberg (eds) (1991) *English Corpus Linguistics.* London: Longman.

Alessi, G. (2005) 'The Use of Metadiscourse in EAP Presentations by Native Italian Speakers' in Bamford, J. and M. Bondi (eds), pp. 179–189.

Altenberg, B and M. Tapper (1998) 'The use of adverbial connectors in advanced Swedish learners' written English' in Granger, S. (ed), pp. 80–93.

Álvarez de Mon y Rego, I. (2001) 'Encapsulation and prospection in written scientific English' *Estudios de la Universidad Complutense* 9: 81–101.

Álvarez de Mon y Rego, I. (2006) 'A contrastive study of encapsulation and prospection in written scientific text' in Flowerdew, J., and M. Gotti (eds), pp. 21–40.

Arendholz, J. (2013) *(In)Appropriate Online Behavior. A Pragmatic Analysis of Message Board Relations*. Amsterdam: Benjamins.

Auer, P. (2005) 'Delayed self-repairs as a structuring device for complex turns in conversation' in Hakulin, A. and M. Selting (eds), pp. 75–102.

Austin, J.L. (1962) *How to Do Things with Words*. Oxford: Oxford University Press.

Baker, M., G. Francis and E. Tognini-Bonelli (eds) (1993) *Text and Technology: In Honour of John Sinclair*. Philadelphia/Amsterdam: John Benjamins Publishing Company.

Bamford, J. and M. Bondi (eds) (2005) *Dialogue within discourse communities*, Niemeyer: Tubingen.

Baron, N. (2009) 'The Myth of Impoverished Signal: Dispelling the Spoken-Language Fallacy for Emoticons in Online Communication' in Vincent, J. and L Fortunati (eds), pp. 107–136.

Barton, E.L. (1995) 'Contrastive and Non-Contrastive Connectives: Metadiscourse Functions in Argumentation' *Written Communication* 12/2: 219–239.

de Beaugrande, R. (1984) *Text Production*. Norwood NJ: Ablex.

Beauvais, P.L. (1989) 'A Speech Act Theory of Metadiscourse' *Written Communication* 6/1: 11–30.

Berry, M. (1981) 'Systemic linguistics and discourse analysis: a multilayered approach to exchange structure' in Coulthard, M. and M. Montgomery (eds), pp. 120–45.

Berwick, R., S. Abney and C. Tenny (eds) (1991) *Principle-Based Parsing*. Dordrecht: Luwer.

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

Blakemore, D. (2004) 'Discourse Markers' in Horn, L.R. and G. Ward (eds), pp. 221–240.

Bolden, G.B. (2009) 'Implementing incipient actions: The discourse marker 'so' in English conversation' *Journal of Pragmatics* 41: 974–998.

Bondi, M. (2005) 'Metadiscourse Practices in Academic Discourse: Variation across Genres and Disciplines' in Bamford, J. and M. Bondi (eds), pp. 3–29.

Brazil, D. (1995) *A Grammar of Speech*. Oxford: Oxford University Press.

Bretones Callejas, C. M., J.F. Fernández Sánchez, J.R. Ibáñez Ibáñez, Mª. E. García Sánchez, Mª E. Cortés de los Ríos, S. Salaberri Ramiro, Mª S. Cruz Martínez, N. Perdú Honeyman and B. Cantizano Márquez (eds) (2009) *Applied Linguistics Now: Understanding Language and Mind / La Lingüística Aplicada Hoy: Comprendiendo el Lenguaje y la Mente*. Almería: Universidad de Almería.

Brinton, L.J. (1996) *Pragmatic Markers in English: Grammaticalization and Discourse Functions.* Berlin/ New York: Mouton de Gryter.

Brown, J.D. (1988) *Understanding Research in Second Language Learning*. Cambridge: Cambridge University Press.

Bublitz, W. and A. Hübler (eds) (2007) *Metapragmatics in Use*. Amsterdam/Philadelphia: John Benjamins.

Bublitz, W. and A. Hübler (2007) 'Introducing metapragmatics in use' in Bublitz, W. and A. Hübler (eds), pp. 1–28.

Bunton, D. (1999) 'The use of higher level metatext in PhD theses' *English for Specific Purposes* 18: S41–S56.

Burton, D. (1980) *Dialogue and Discourse*. London: Routledge & Kegan Paul.

Burton. D. (1981) 'Analyzing Spoken Discourse' in Coulthard, M. and M. Montgomery (eds), pp. 61–81.

Butler, C.S. (2008) 'The Subjectivity of *basically* in British English – a corpus-based study' in Romero-Trillo, J. (ed), pp. 37–62.

Bybee, J. (1998) 'The Emergent Lexicon' *CLS 34: The Panels*: 421–435.

Bybee, J and P. Hopper (2001) 'Introduction to frequency and the emergence of linguistic structure' in Bybee, J. and P. Hopper (eds), pp. 1–24.

Bybee, J. and P. Hopper (eds) (2001) *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamin Publishing Company.

Carretero, M., L. Hidalgo Downing, J. Lavid, E. Martinez Caro, J. Neff, S. Perez de Ayala and E. Sanchez-Pardo (eds) (2006) *A Pleasure of Life in Words: A Festschrift for Angela Downing*. Madrid: Universidad Complutense Madrid.

Carey, R. (2011) 'Interruption and Uncooperativeness in Academic ELF Group Work: An Application of Linear Unit Grammar' unpublished Masters Thesis, Department of Modern Languages, University of Helsinki. Accessible: https://helda.helsinki.fi/bitstream/handle/10138/28956/interrup.pdf?sequence=1 . Accessed 19/08/2013.

Carey, R. (2013) 'On the other side: formulaic organizing chunks in spoken and written academic EFL' *Journal of English as a Lingua Franca* 2/2:207–228.

Chalker, S. (1996) *Collins COBUILD English Guides 9: Linking Words*. London: HarperCollins Publishers Limited.

Charles, M. (2003) ''This mystery…': a corpus-based study of the use of nouns to construct stance in theses from two contrasting disciplines' *Journal of English for Academic Purposes* 2/4:313–326.

Chávez Muñoz, M. (2013) 'The "I" in interaction: authorial presence in academic writing' *Revista de Lingüística y Lenguas Aplicadas* Vol. 8, pp. 49–58.

Cheng, W. (2007) 'Concgramming: A Corpus-Driven Approach to Learning the Phraseology of Discipline-Specific Texts' *CORELL: Computer Resources for Language Learning* 1: 22–35.

Cheng, W., C. Greaves, J. McH. Sinclair and M. Warren (2009) 'Uncovering the Extent of the Phraseological Tendency: Towards a Systematic Analysis of Concgrams' *Applied Linguistics* 30/2: 236–252.

Cheng, W, C. Greaves and M. Warren (2008) *A corpus-driven study of discourse intonation: the Hong Kong corpus of spoken English (prosodic).* Studies in Corpus Linguistics, vol. 32. Amsterdam: John Benjamins.

Cheng, X. and M. Steffensen (1996) 'Metadiscourse: a technique for improving student writing' *Research in the Teaching of English* 30(2): 149–81.

Claridge, C. (2007) 'Constructing a Corpus from the Web: Message Boards' in Hundt, M. et al (eds), pp. 87–108.

Coffin, C., A. Hewings and S. North (2012) 'Arguing as an academic purpose: The role of asynchronous conferencing in supporting argumentative dialogue in school and university' *Journal of English for Academic Purposes* 11(1): 38–51.

Collins Cobuild Advanced Learner's English Dictionary, Fourth Edition (2003) London: HarperCollins.

Collot, M. and N. Belmore (1996) 'Electronic Language: A new variety of English' in Herring, S. (ed), pp. 13-28.

Connor, U., E. Nagelhout and W.V. Rozycki (eds) (2008) *Contrastive Rhetoric: Reaching to intercultural rhetoric*. Amsterdam/Philadelphia: John Benjamins.

Cortese, G. and P. Riley (eds) *Domain-Specific English: Textual Practices Across Communities and Classrooms*. Frankfurt: Peter Lang: 115–140.

Coulthard, M. (ed) (1992) *Advances in Spoken Discourse Analysis*. London: Routledge.

Coulthard, M (ed) (1994) *Advances in Written Text Analysis*. London: Routledge.

Coulthard, M and D. Brazil (1992) 'Exchange structure' in Coulthard, M (ed), pp. 50–78.

Coulthard, M. and M. Montgomery (eds) (1981) *Studies in Discourse Analysis*. London: Routledge & Kegan Paul.

Coulthard, M., M. Montgomery and D. Brazil (1981) 'Developing a description of spoken discourse' in Coulthard M. and M. Montgomery (eds), pp. 1–50.

Crawford Camiciottoli, B. (2003) 'Metadiscourse and ESP reading comprehension: An explanatory study' *Reading in a Foreign Language* 15/1: 28–44.

Crawford Camiciottoli, B. (2007) The *Language of Business Studies Lectures. A corpus-assisted analysis.* Amsterdam/Philadelphia: John Benjamins.

Crismore, A. (1983) *Metadiscourse: what it is and how it is used in school and non-school science texts (Tech. Rep. No. 273)*. Champaign: University of Illinois, Center for the Study of Reading.

Crismore. A. (1989) *Talking with Readers: Metadiscourse as Rhetorical Act.* New York*:* Peter Lang.

Crismore, A., R. Markkanen and M.S. Steffensen (1993) 'Metadiscourse in Persuasive Writing' *Written Communication* 10/1: 39–71.

Crismore, A. and R. Farnsworth (1990) 'Metadiscourse in popular and professional science discourse' in Nash (ed), pp. 118–36.

Crombie, W. (1985) *Process and Relation in Discourse and Language Learning*. Oxford: Oxford University Press.

Crompton, P. (2006) 'The effect of the position on the Discourse Scope of Adverbials' *Text and Talk* 26/3: 245–279.

Crystal, D. (2001) *Language and the Internet*. Cambridge: Cambridge University Press.

Dafouz-Milne, E. (2008) 'The pragmatic role of textual and interpersonal metadiscourse markers in the construction and attainment of persuasion: a cross-linguistic study of newspaper discourse' *Journal of Pragmatics* 40 (1): 95–113.

Dahl, T. (2004) 'Textual metadiscourse in research articles: a marker of national culture or academic discipline' *Journal of Pragmatics* 36(10): 1807–25.

Dillon, G.L. (1981) *Constructing Texts, Elements of a Theory of Composition and Style*. Indiana: Indiana University Press.

Dunning, T. (1993) 'Accurate methods for the statistics of surprise and coincidence' *Computational Linguistics* 19/1: 61–74.

Eggins, S. (2004) *An Introduction to Systemic Functional Linguistics.* London: Continuum International Publishing Group.

Eggins, S and D. Slade (1997) *Analysing casual conversation*. London: Continuum International Publishing Group.

Erman, B. (2001) 'Pragmatic markers revisited with a focus on *you know* in adult and adolescent talk' *Journal of Pragmatics* 33: 1337–1359.

Farooq, M.U. (2000) 'Examining a Male Teacher's Attention in a Mixed-sex EFL Japanese High School Classroom Based on the Sinclair-Coulthard model', unpublished MA dissertation, University of Birmingham.

Ferrara, K., H. Brunner and G. Whittemore (1991) 'Interactive written discourse as an emergent register' *Written communication* 8/1: 8–34.

Fisiak, J. (ed) (1993) *Papers and Studies in Contrastive Linguistics.* Poznan: Mickiewicz University.

Fløttum, K., T. Dahl and T. Kinn (2006) *Academic Voices: Across languages and disciplines*. John Benjamins Publishing Company: Amsterdam/Philadelphia.

Flowerdrew, J. (2003) 'Signalling Nouns in Discourse' *English for Specific Purposes* 22: 329–46.

Flowerdew, J. and M. Gotti (eds) (2006) *Studies in Specialized Discourse*. Bern: Peter Lang.

Francis, G. (1986) *Anaphoric nouns. Discourse analysis monographs No. 11*. English Language Research, University of Birmingham.

Francis, G. (1994) 'Labelling discourse: an aspect of nominal-group lexical cohesion' in Coulthard, M. (ed), pp. 83–101.

Francis. G and S. Hunston (1992) 'Analysing everyday conversation' in Coulthard, M (ed), pp. 123–161.

Francis, G., S. Hunston and E. Manning (1996) *Collins COBUILD Grammar Patterns 1: Verbs*. London: HarperCollins.

Francis, G., S. Hunston and E. Manning (1998) *Collins COBUILD Grammar Patterns 2: Nouns and Adjectives*. London: HarperCollins.

Fraser, B. (1990) 'An approach to discourse markers' *Journal of Pragmatics* 14: 383–395.

Fraser, B. (1999) 'What are discourse markers?' *Journal of Pragmatics* 31: 931–952.

Fuertes-Olivera, P.A., M. Velasco-Sacristán, A. Arribas-Baño and E. Samaniego-Fernández (2001) 'Persuasion and advertising English: Metadiscourse in slogans and headlines' *Journal of Pragmatics* 33: 1291–1307.

Gerbig, A and O. Mason (eds) (2008) *Language, People, Numbers: Corpus Linguistics and Society*. Amsterdam: Rodopi.

Ghaddessy, M. (ed) (1999) *Text and context in functional linguistics*. Amsterdam: Benjamins.

Gibbs, R.W. (2000) 'Irony in Talk Among Friends' *Metaphor and Symbol* 15/1–2: 5–27.

Granger, S. (ed) (1998) *Learner English on Computer*. Harlow: Addison Wesley Longman Limited.

Greenberg, J.H. (ed) (1963) *Universals of Language*. Cambridge, Massachusetts: The M.I.T. Press.

Hakulin, A. and M. Selting (eds) (2005) *Syntax and Lexis in Conversation: Studies on the use of linguistic resources in talk-in-interaction*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Halliday, M.A.K. (1961) 'Categories of the theory of grammar' *Word* 17/3: 241–292.

Halliday, M.A.K. (1973) *Explorations in the Functions of Language*. London: Arnold.

Halliday, M.A.K. (1994) *An Introduction to Functional Grammar (Second Edition).* London: Arnold.

Halliday M.A.K. and R. Hasan. (1976) *Cohesion in English*. London: Longman.

Harris, Z.S. (1970) *Papers in Structural and Transformational Linguistics*, Dordrecht: D. Reidel Publishing Company.

Harrison, S. (1998) 'E-mail discussions as conversation: moves and acts in a sample from a listserv discussion' *Linguistik online, Computer-vermittelte Kommunikation Computer-mediated communication* 1/98. Accessible: http://www.linguistik-online.de/harrison.htm. Accessed on 19/08/2013.

Hasselgård, H and S. Oksefjell (eds) (1999) *Out of Corpora: Studies in Honour of Stig Johansson*. Amsterdam/Atlanta, GA: Rodopi.

Hatter, J.M. (2002) 'The Application of Exchange Theory to Internet Relay Chat' unpublished MA Dissertation, University of Birmingham.

Hazadiah, M.D. (1993) 'Topic as a Dynamic Element in Spoken Discourse' in Baker, M., G. Francis and E. Tognini-Bonelli (eds), pp. 55–74.

Heino, A., E. Tervonen and J. Tommola (2002) 'Metadiscourse in Academic Conference Presentations' in Ventola, E., C. Shalom and S. Thompson (eds), pp. 127–146.

Herriman, J. (2007) '"I'm overstating my case": Overt Authorial Presence in English Argumentative Texts by Students and Professional Writers' *Nordic Journal of English Studies* 6/1. Accessible: http://ojs.ub.gu.se/ojs/index.php/njes/issue/view/6. Accessed: 25/09/2013.

Herring, S. C. (ed) (1996) *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*. Amsterdam/Philadelphia: John Benjamins.

Herring, S. C. (1999) 'Interactional coherence in CMC' *Journal of Computer-Mediated Communication, 4* (4). Accessible: http://jcmc.indiana.edu/vol4/issue4/herring.html Accessed: 19/08/2013.

Hewings, A. and C. Coffin (2007) 'Writing in multiparty computer conferences and single authored assignments: Exploring the role of writer as thinker' *Journal of English for Academic Purposes* 6: 126–142.

Ho, M.L.C. (2002) 'Online communication: A study of the construction of discourse and community in an electronic discussion forum' unpublished PhD thesis, University of Birmingham.

Hockett, C. (1963) 'The Problem of Universals in Language' in Greenberg (ed), pp. 1–22.

Hoey, M. (1983) *On the Surface of Discourse*. London: Allen & Unwin.

Hoey, M. (2000) 'Persuasive Rhetoric in Linguistics: A Stylistic Study of some Features of the Language of Noam Chomsky' in Hunston, S. and G. Thompson (eds), pp. 28–38.

Hoey, M. (2001) *Textual interaction: an introduction to written discourse analysi*s. Routledge: London and New York.

Horn, L.R. and G. Ward (eds) (2004) *The Handbook of Pragmatics*. Oxford: Blackwell Publishing Ltd.

Hopper, P.J. and E.C. Traugott (1993) *Grammaticalization.* Cambridge, UK: Cambridge University Press.

Hornby, N. (1995) *High Fidelity.* London: Penguin Book Ltd.

Hundt, M., N. Nesselhauf and C. Biewer (eds) (2007) *Corpus Linguistics and the Web*. Amsterdam/New York: Rodopi.

Hunston, S. (2000) 'Evaluation and the planes of discourse: status and value in persuasive texts' in Hunston, S. and G. Thompson (eds), pp. 176–207.

Hunston, S. and G. Thomson (eds) (2000) *Evaluation in Text: authorial stance and the construction of discourse.* Oxford: Oxford University Press.

Hunston, S. and G. Francis (2000) *Pattern Grammar*. Amsterdam: John Benjamins.

Hyland, K. (1998a) *Hedging in Scientific Research Articles*. Amsterdam: John Benjamins.

Hyland, K. (1998b) 'Persuasion and context: The pragmatics of academic discourse' *Journal of Pragmatics* 30/4: 437–455.

Hyland, K. (1998c) 'Boosting, hedging and the negotiation of academic knowledge' *TEXT* 18/3: 349–82.

Hyland, K. (1998d) 'Exploring Corporate Rhetoric', *Journal of Business Communication* 35/2: 224–245.

Hyland, K. (1999) 'Talking to Students: metadiscourse in introductory textbooks' *English for Specific Purposes* 18(1): 3–26.

Hyland, K. (2000) *Disciplinary Discourse: Social Interactions in Academic Writing* Harlow: Pearson Education Ltd.

Hyland, K. (2005) *Metadiscourse: Exploring Interaction in Writing*. London: Continuum.

Hyland, K. (2010) 'Metadiscourse: Mapping Interactions in Academic Writing' *Nordic Journal of English Studies* 9/2: 125–143.

Hyland, K. and P. Tse (2004) 'Metadiscourse in academic writing: a reappraisal' *Applied Linguistics* 25(2): 156–77.

Hyland, K. and P. Tse (2005) 'Evaluative *that* constructions: Signalling stance in research abstracts' *Functions of Language* 12/1: 39–63.

Ifantidou, E. (2005) 'The semantics and pragmatics of metadiscourse' *Journal of Pragmatics* 37: 1325–1353.

Ilie, C. (2003) 'Discourse and metadiscourse in parliamentary debates' *Journal of Language and Politics* 2/1: 71–92.

Intaraprawat, P. and M. Steffensen (1995) 'The use of metadiscourse in good and poor ESL essays' *Journal of Second Language Writing* 4/3: 253–72.

Ivanič, R. (1991) 'Nouns in Search of a Context: A Study of Nouns with both Open- and Closed-System Characteristics' *International Review of Applied Linguistics* 29/2: 93–114.

Jakobson, R. (1980) *The Framework of Language*. Michigan: Michigan Studies in the Humanities.

Jakobson, R. (1990) *On Language: Roman Jakobson*. L.R. Waugh and M. Monville-Burston (eds). Cambridge, MA and London: Harvard University Press.

Johnson, S. (2008-12-28) 'Rap music originated in medieval Scottish pubs, claims American professor' in *The Telegraph*, Accessible:http://www.telegraph.co.uk/culture/music/3998862/Rap-music-originated-in-medieval-Scottish-pubs-claims-American-professor.html. Accessed 19/08/2013.

Jones, C. and E. Ventola (eds) (2008) *New Developments in the Study of Ideational Meaning. From Language to Multimodality*. London: Equinox.

Jones, R.H. and C.A. Hafner (2012) *Understanding Digital Literacies*: *A Practical Introduction*. London: Routledge.

Khabbazi-Oskouei, L. (2013) 'Propositional or non-propositional, that is the question: A new approach to analyzing 'interpersonal metadiscourse in editorials' *Journal of Pragmatics* 47/1: 93–107.

Kohonen, V. and N.E. Enkvist (eds) (1978) *Text Linguistics Cognitive Learning and Language Teaching*. Turku, Finland: Finnish Association of Applied Linguistics.

Kuteeva. M. (2011) 'Wikis and academic writing: Changing the writer-reader relationship' *English for Specific Purposes* 30: 44–57.

Lam, P.W.Y. (2010) 'Discourse Particles in Corpus Data and Textbooks: The Case of Well' *Applied Linguistics* 31/2: 260–281.

Landis J.R. and G.G. Koch (1977) 'The measurement of observer agreement for categorical data' *Biometrics* 33:159-174.

Lautamatti, L. 1978 'Observations on the development of topic in simplified discourse' in Kohonen, V. and N.E. Enkvist (eds), pp. 71–104.

Le, E. (2004) 'Active participation within written argumentation: metadiscourse and editorialist's authority' *Journal of Pragmatics* 36: 687–714.

Leech, G. (1999) 'The Distribution and Function of Vocatives in American and British Conversation' in Hasselgård, H. and S. Oksefjell (eds), pp. 107–120.

Leech, G., M. Deuchar and R. Hoogenraad (1982). *English Grammar for Today: A New Introduction*. London: MacMillan.

Leech, G., P. Rayson, A, Wilson (2001) *Word Frequencies in Written and Spoken English: based on the British National Corpus*. London: Longman.

Leistyna, P. and C.F. Meyer (eds) (2003) *Corpus Analysis. Language Structure and Language Use.* Amsterdam: Rodopi.

Levelt, W. (1989) *Speaking*. Cambridge MA: MIT Press.

Levinson, S. (1983) *Pragmatics*. Cambridge: Cambridge University Press.

Lewin, B.A. and Y. Donner (2002) 'Communication in internet message boards' *English Today* 18/3: 29–37.

Li, T., and S. Wharton (2012) 'Metadiscourse repertoire of L1 Mandarin undergraduates writing in English: A cross-contextual, cross-disciplinary study' *Journal of English for Academic Purposes* 11/4: 345–356.

Longo, B. (1994) 'Current research in technical communication: The role of metadiscourse in persuasion' *Technical Communication* 4/2: 348–352.

Lozano, C. (2009) 'Pronominal deficits at the interface: New data from the CEDEL2 corpus' in Bretones Callejas, C. M. et al (eds), pp. 213–227.

Lucy, J.A. (ed) (1993) *Reflexive language: Reported speech and metapragmatics*. Cambridge: Cambridge University Press.

Lucy, J.A. (1993) 'Reflexive language and the human disciplines' in Lucy, J.A. (ed), pp. 9–33.

Lyons, J. (1977) *Semantics Vol. I*, New York: Cambridge University Press.

McEnery, T. and A. Wilson (2001) *Corpus Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

Mahlberg, M. and M.B. O'Donnell (2008) 'A *fresh* view of the structure of hard news stories', in Neumann, S. and E Steiener (eds). Accessible: http://scidok.sulb.uni-saarland.de/volltexte/2008/1700/ . Accessed 18/08/2013.

Mann, W. and S.A. Thompson (1986) 'Relational Propositions in Discourse' *Discourse Processes* 9/1: 57–90.

Mann, W.C. and S.A. Thompson (1988) 'Rhetorical Structure Theory: Toward a Functional Theory of Text Organization' *Text* 8/3: 243–81.

Mao, L.R. (1993) 'I Conclude Not: Toward a Pragmatic Account of Metadiscourse' *Rhetoric Review* 2 (Spring): 265–289.

Marcoccia, M. (2004) 'On-line Polylogues: Conversation Structure and Participation Frameworks in Internet Newsgroups' *Journal of Pragmatics* 36: 115–145.

Markham, A and E. Buchanan (2012) *Ethical decision-making and Internet research: recommendations from the AoIR ethics working committee*. Accessible: www.aoir.org/reports/ethics.pdf. Accessed: 19/08/2013.

Markkanen, R., M.S. Steffensen and A. Crismore (1993) 'A quantitative contrastive study of metadiscourse: problems in the design and analysis of the data' in Fisiak, J. (ed.), pp. 137–51.

Martín-Úriz, A. and S. Murcia-Bielsa (2008) 'Reports in Spanish writing: generic features and text quality' in Wu, C. et al (eds), pp. 72–78.

Martín-Úriz, A., R. Whittaker, S. Murcia and K. Vidal (2008) 'The expression of experiential meaning in EFL students' texts: an analysis of secondary school recounts' in Jones, C. and E. Ventola (eds), pp. 209–227.

Mason. O. (2007) 'Review of Linear unit Grammar: Integrating speech and writing' *TESEL–EJ* 11/2. Accessible: http://www.tesl-ej.org/wordpress/. Accessed 19/08/2013.

Mason. O. (2008) 'Stringing together a sentence: linearity and the lexis-syntax interface' in Gerbig, A. and O. Mason (eds), pp. 231–248.

Mauranen, A. (1993a) *Cultural Differences in Academic Rhetoric: A Textlinguistic Study*. Frankfurt an Maim: Peter Lang.

Mauranen, A. (1993b) 'Contrastive ESP Rhetoric: Metatext in Finnish-English Economics Texts' *English for Specific Purposes* 12: 3–22.

Mauranen, A. (1996) 'Discourse Competence–Evidence from Thematic Development in Native and Non-native Texts' in Ventola, E and A. Mauranen (eds), pp. 231–235.

Mauranen, A. (2001) 'Reflexive academic talk: observations from MICASE' in Simpson, R. and J.M. Swales (eds), pp. 165–178.

Mauranen, A. (2002a) '"One Thing I'd Like to Clarify…". Observations of Academic Speaking' Helsinki English Studies 2002, Vol. 2. Accessible: http://blogs.helsinki.fi/hes-eng/files/2011/03/HES_Vol2_Mauranen.pdf. Accessed: 19/08/2013.

Mauranen, A. (2002b) '"A good question": expressing evaluation in academic speech' in Cortese, G. and P. Riley (eds), pp. 115–140.

Mauranen, A. (2003) '"But here's a flawed argument": socialization into and through metadiscourse' in Leistyna, P. and C.F. Meyer (eds), pp. 19–34.

Mauranen, A. (2007) 'Discourse Reflexivity and International Speakers– How is it used in English as a Lingua Franca?' *Jesik in Slovstvo* 52 /3–4, S1–S19.

Mauranen, A. (2009a) 'Spoken rhetoric: How do natives and non-natives fare?' in Suomela-Salmi, E. and F. Dervin (eds), pp. 199–218.

Mauranen, A. (2009b) 'Chunking in ELF: Expressions for managing interaction' *Journal of Intercultural Pragmatics* 6/2: 217–233.

Mauranen A. (2010) 'Discourse Reflexivity – A Discourse Universal? The Case of ELF' *Nordic Journal of English Studies* 9/2: 13–40.

Mauranen, A. (2012) *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge: Cambridge University Press.

Meyer, B.J.F. (1975) *The Organization of Prose and its Effects on Memory*. Amsterdam: North-Holland Publishing Company.

Milizia D. and C. Spinzi. (2008) 'The terroridiom principle between spoken and written discourse' Special issue of *International Journal of Corpus Linguistics* 13/3: 322–350.

Miller, G. (1956) 'The magical number seven, plus or minus two: Some limits on our capacity for processing information' *Psychological Review* 63: 81–97.

Monbiot, G. (2010) 'These astroturf libertarians are the real threat to internet democracy' *The Guardian*, Monday 13 December 2010. Available online at: http://www.guardian.co.uk/commentisfree/libertycentral/2010/dec/13/astroturf-libertarians-internet-democracy?showallcomments=true#comment-fold. Accessed 19/08/2013.

Moon, R. (1994) 'The analysis of fixed expressions in text' in Coulthard (ed), pp. 117–135.

Moor, P.J., A. Heuvelman and R. Verleur (2010) 'Flaming on YouTube' *Computers and Human Behavior* 26(6): 1536–1546.

Moore, N. (2011) 'Modelling the Flow of Discourse in a Corpus of Written Academic English' *Proceedings of Corpus Linguistics 2011 – Discourse and Corpus Linguistics*. Accessible : http://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2011/Paper-11.pdf. Accessed: 14/09/2013.

Moreno, A.I. (2003) 'The Role of Cohesive Devises as Textual Constraints on Relevance: A discourse-as-process View' *International Journal of English Studies* 3/1: 111–165.

Moreno, A.I. (2004) 'Retrospective labelling in premise–conclusion metatext: an English – Spanish contrastive study of research articles on business and economics' *Journal of English for Academic Purposes* 3: 321–339.

Moreno, A. (2006) 'Bridging the gap between coherence and cohesion: which cohesive devices are really textual?' in Carretero, M et al (eds), pp. 673–708.

Mukherjee, J. (2001) *Form and Function of Parasyntactic Presentation Structures. A corpus-based study of talk units in spoken English*. Amsterdam: Rodopi.

Müller. S. (2005) *Discourse Markers in Native and Non-native English Discourse*. Amsterdam: John Benjamins.

Muntigl, P. (2007) 'A metapragmatic examination of therapist reformulations' in Bublitz, W. and A. Hübler (eds), pp. 235–262.

Mur-Dueñas, M.P. (2004) 'Analysing stance in American and Spanish business Management RAs: The Case of Sentence-Initial 'Retrospective Labels'' *International Journal of English Studies* 4: 137–154

Myers, G. (2010) *Discourse of Blogs and Wikis*. London: Continuum.

Nash, W. (ed) (1990) *The Writing Scholar: Studies in Academic Discourse.* Newbury Park, CA: Sage.

Nash, W. (1992) *The Uncommon Tongue: the uses and resources of English*. London: Routledge.

Neumann, S. and E Steiener (eds) *Online Proceedings of the 19th European Systemic Functional Linguistics Conference and Workshop*. Saarbrücken, 23-25 July 2007 Accessible: http://scidok.sulb.uni-saarland.de/volltexte/2008/1700/. Accessed 18/08/2013.

O'Donnell, M. (1990) 'A Dynamic Model of Exchange' *Word* 41/3: 293–328.

O'Donnell, M. and P. Sefton (1995) 'Modelling Telephonic Interaction: A Dynamic Approach' *Journal of Applied Linguistics* 10/1: 63–78.

O'Donnell, M. (1999) 'Context in Dynamic Modelling' in Ghadessy, M. (ed.), pp. 63–99.

O'Donnell, M. (2009) 'The UAM CorpusTool: Software for corpus annotation and exploration' in Bretones Callejas, C.M. et al (eds), pp. 1433–1447.

O'Donnell, M.B., M. Scott, M. Mahlberg, and M. Hoey (2012) 'Exploring text-initial words, clusters and concgrams in a newspaper corpus' *Corpus Linguistics and Linguistic Theory* 8–1: 73–101.

O'Grady, G. (2010) *A Grammar of Spoken English Discourse.* London: Continuum International Publishing Group.

Paltridge, B. (2006) *Discourse Analysis*. London: Continuum.

Parsons, B. (2011) 'An Early Flyting in Hary's Wallace' *Marginalia* 13: *Taste* Accessible: http://www.marginalia.co.uk/journal/11taste/11tasteparsons.pdf. Accessed 19/08/2013.

Pérez-Llantada C. (2010) 'The Discourse Functions of Metadiscourse in Published Academic Writing: Issues of Culture and Language' *Nordic Journal of English Studies* 9/2: 41–68.

Pihlaja, S. (2011) 'Cops, popes, and garbage collectors: Metaphor and antagonism in an atheist/Christian YouTube video thread' *Language@Internet*, 8, article 1. Accessible: http://oro.open.ac.uk/29116/1/www_languageatinternet_de_articles_2011_pihlaja.pdf. Accessed 19/08/2013.

Purnelle G., C. Fairon, A. Dister (eds) (2004) *Le poids des mots: Proceedings of the 7th International Conference on Statistical analysis of textual data (JADT 2004), March 10–12, 2004 (Vol. 2)*: Louvain-la-Neuve, Belgium, Presses universitaires de Louvain.

Rahman, M. (2004) 'Aiding the Reader: The use of Metalinguistic Devices in Scientific Discourse' *Nottingham Linguistic Circular*, 18.

Rayson P., D. Berridge and B. Francis (2004) 'Extending the Cochran rule for the comparison of word frequencies between corpora' in Purnelle G., C. Fairon and A. Dister (eds), pp. 926 - 936.

Rayson, P. and R. Garside (2000) 'Comparing corpora using frequency profiling' in *Proceedings of the workshop on Comparing Corpora, held in conjunction with the 38th annual meeting of the Association for Computational Linguistics (ACL 2000). 1–8 October 2000, Hong Kong*, pp. 1 – 6.

Accessible: http://www.comp.lancs.ac.uk/~paul/publications/rg_acl(2000)pdf. Accessed 19/08/2013.

P. Rayson, A. Wilson, T. McEnery, A. Hardie and S. Khoja (eds) (2001) *UCREL Technical Paper number 13. Special issue. Proceedings of the Corpus Linguistics 2001 conference*. Lancaster University, 29 March– 2 April (2001).

Renouf, A. and J. McH. Sinclair (1991) 'Collocational frameworks in English' in Aijmer, K. and B. Altenberg (eds), pp. 128–44.

Ribera, J. (2007) 'Text Deixis in Narrative Sequences' *International Journal of English Studies* 7/1: 149–168.

Richards, J., J. Platt and H. Platt (1992) *Longman Dictionary of Language Teaching & Applied Linguistics*. Harlow: Longman.

Romero-Trillo, J. (ed) (2008) *Pragmatics and Corpus Linguistics: A mutualistic entente*. Berlin: Mouton de Gruyter.

Saussure, F. de (1959/1966). *Course in General Linguistics*. C. Bally and A. Sechehaye (eds). New York: McGraw-Hill.

Schiffrin, D. (1980) 'Metatalk: organizational and evaluative brackets in discourse' *Sociological inquiry: Language and Social Interaction* 50: 199–236.

Schiffrin, D. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.

Schmid, H.J. (2000). *English Abstract Nouns as Conceptual Shells. From Corpus to Cognition*. Berlin: Mouton de Gruyter.

Scott, M. (1996) *Wordsmith Tools.* Oxford: Oxford University Press.

Scott, M. (1997) 'PC analysis of key words – and key key words' *System* 25(2): 233–45.

Scott, M. (2001) 'Mapping key words to problem and solution' in Scott, M. and G. Thomson, (eds), pp. 109–127.

Scott, M. and G. Thompson (eds) (2001) *Patterns of Text: in honour of Michael Hoey*. Amsterdam: Benjamins.

Searle, J.R. (1969) *Speech Acts: an essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Shapen, T. and J.M. Williams (eds). (1981) *Style and variables in English*. Cambridge, MA: Winthorp.

Sheldon, E. (2009) 'From one I to another: Discursive construction of self-representation in English and Castilian Spanish research articles' *English for Specific Purposes* 28/4: 251–261.

Siepman, D. (2005) *Discourse Markers Across Languages: A Contrastive Study of Second-Level Discourse Markers in Native and Non-Native Text with Implications for General and Pedagogic Lexicography.* Oxon: Routledge.

Simpson, R. and J.M. Swales (eds) (2001) *Corpus Linguistics in North America*, Ann Arbor: University of Michigan Press.

Sinclair, J. McH. (1991) *Corpus, Concordance and Collocation*. Oxford: Oxford University Press.

Sinclair, J. (1992) 'Priorities in discourse analysis' in Coulthard, M. (ed), pp. 79-88.

Sinclair, J. McH. (1992/2004a) 'Trust the text' in Sinclair, J. McH. and R. Carter (eds), pp. 9–23.

Sinclair, J. McH. (1996/2004b) 'The search for units of meaning' in Sinclair, J. McH. and R. Carter (eds), pp. 24–48.

Sinclair, J. McH. (1981/2004c) 'Planes of discourse' in Sinclair, J. McH. and R. Carter (eds), pp. 51–66.

Sinclair, J. McH. (1985/2004d) 'On the integration of linguistic description' in Sinclair, J. McH. and R. Carter (eds), pp. 67–81.

Sinclair, J. McH. (1993/2004e) 'Written discourse structure' in Sinclair, J. McH. and R. Carter (eds), pp. 82–101.

Sinclair, J. McH. (1998/2004f) 'The lexical item' in Sinclair, J. McH. and R. Carter (eds), pp. 131–148.

Sinclair, J. McH. (2003/2004g) 'Lexical grammar' in Sinclair, J. McH. and R. Carter (eds), pp. 164–176.

Sinclair, J. (2005) 'Language as a string of beads: Discourse and the M-word' in Tognini-Bonelli and G. Dell Lungo Camiciotti (eds), pp. 163–168.

Sinclair, J. McH. (2007) 'Collocation reviewed' Manuscript Tuscan Word Centre, Italy.

Sinclair, J. McH. and R. Carter (eds) (2004) *Trust the Text: language, corpus and discourse*. London: Routledge Taylor and Francis Group.

Sinclair, J McH. and R.M. Coulthard (1975) *Towards an Analysis of Discourse*. London: Oxford University Press.

Sinclair, J. McH. and A. Mauranen (2006) *Linear Unit Grammar: Integrating speech and writing*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Skulstad, A.S. (2005) 'The use of metadiscourse in introductory sections of a new genre' *International Journal of Applied Linguistics* 15/1: 71–86.

Stenström, A-B. (1990) 'What is the Role of Discourse signals in Sentence Grammar?' in Aarts, J. and W. Meijs (eds), pp. 213–229.

Stenström, A-B. (1994) *An introduction to Spoken Interaction*. Harlow: Addison Wesley Longman Limited.

Stubbs, M. (1981) 'Motivating analyses of exchange' in Coulthard, M. and M. Montgomery (eds), pp. 107–119.

Stubbs, M. (1983) *Discourse Analysis: The Sociolinguistic Analysis of Natural Language.* Chicago: University of Chicago Press.

Stubbs, M. (1996) *Text and Corpus Analysis.* Oxford: Blackwell.

Suomela-Salmi, E. and F. Dervin (eds) (2009) *Cross-linguistic and Cross-cultural Perspectives on Academic Discourse*. Amsterdam: John Benjamins.

Swan, M. (2005) *Practical English Usage Third Edition*. Oxford: Oxford University Press.

Tadros, A. (1985) *Prediction in Text*. University of Birmingham: English Language Research.

Tadros, A. (1994) 'Predictive categories in expository text' in Coulthard, M. (ed), pp. 69–82.

Tagliamonte, S. and D. Denis. (2008) 'LOL for Real! Instant Messaging in Toronto Teens' *American Speech* 83:1: 3–34.

Tagg, C. (2012) *The Discourse of Text Messaging*. London: Continuum.

Tang, R., & S. John (1999) 'The 'I' in identity: Exploring writer identity in student academic writing through the first person' *English for Specific Purposes* 18, S23–S39.

Tanskanen, S-K. (2006) *Collaborating Towards Coherence: Lexical Cohesion in English Discourse.* Amsterdam/Philadelphia: John Benjamins.

Tanskanen, S-K. (2007) 'Metapragmatic Utterances in Computer-Mediated Interaction' in Bublitz, W. and A. Hübler (eds), pp. 87–106.

Thompson, G. (2001) 'Interaction in Academic Writing: Learning to Argue with the Reader' *Applied Linguistics* 22/1: 58–78.

Thompson, G. (2012) 'Intersubjectivity in newspaper editorials: construing the reader-in-the-text' *English Text Construction* 5(1): 77–100.

Thompson, G and S. Hunston (2000) 'Evaluation: An Introduction' in Hunston, S. and G. Thomson (eds), pp. 1–27.

Thompson, G. and P. Thetela (1995) 'The sound of one hand clapping: The management of interaction in written discourse' *Text* 15(1): 103–127.

Thompson, G. and J. Zhou (2000) 'Evaluation and Organization in Text: The Structuring Role of Evaluative Disjuncts' in Hunston, S. and G. Thomson (eds), pp. 121–141.

Thompson, S.E. (2003) 'Text-structuring metadiscourse, intonation and the signaling of organization in academic lectures' *English for Academic Purposes* 2: 5–20.

Tognini-Bonelli and G. Dell Lungo Camiciotti (eds) (2005) *Strategies in Academic Discourse*, Amsterdam/Philadelphia: John Benjamins Publishing Company.

Tsui, A. (1994) *English Conversation*. Oxford: Oxford University Press.

Valero-Garces, C. (1996) 'Contrastive ESP Rhetoric: Metatext in Spanish-English Economics Texts' *English for Specific Purposes* 15/4: 279–294.

Van Rooy, B. (2008) 'An alternative interpretation of tense and aspect in Black South African' *World Englishes* 27/3–4: 335–358

Vande Kopple, W.J. (1985) 'Some exploratory discourse on metadiscourse' *College Composition and Communication* 36: 82–93.

Vande Kopple, W.J. (1988) 'Metadiscourse and the Recall of Modality Markers' *Visible Language* Volume XXII, 2/3: 233–272.

Ventola, E. (1987) *The Structure of Social Interaction: a Systemic Approach to the Semiotics of Service Encounters.* London: Pinter.

Ventola, E and A. Mauranen (eds) (1996) *Academic Writing: Intercultural and Textual Issues*. Amsterdam/ Philadelphia: John Benjamin Publishing Co.

Ventola, E., C. Shalom and S. Thompson (eds) (2002) *The Language of Conferencing*, Frankfurt am Main: Peter Lang.

Vincent, J. and L. Fortunati (eds) (2009) *Electronic Emotion: The Mediation of Emotion Via Information and Communication Technologies Interdisciplinary Communication Studies*, *Vol. 3*. Bern, Switzerland: Peter Lang.

Warren, M. (2006) *Features of Naturalness in conversation.* Amsterdam/ Philadelphia: John Benjamin Publishing Co.

Williams, J.W. (1981a) *Style: Ten Lessons in Clarity & Grace.* Glenview, Illinois: Scott, Foresma.

Williams, J.W. (1981b) *Literary Style: The personal voice* in Shapen, T. and J.M. Williams (eds), pp. 116–216.

Winter, E.O. (1977) 'A clause-relational approach to English texts: a study of some predictive lexical items in written discourse' *Instructional Science* (special issue) 6/1: 1–92.

Winter, E. (1994) 'Clause relations as information structure: two basic text structures in English' in Coulthard, M. (ed), pp. 46–58.

Wu, C., C.M.I.M Matthiessen and M. Herke (eds) (2008) *Proceedings of ISFC 35: Voices Around the World*. The 35th ISFC Organizing Committee: Sydney.

Yamasaki N. (2008) 'Collocations and colligations associated with discourse functions of unspecific anaphoric nouns' *International Journal of Corpus Linguistics* 13(1): 75–98.

Zappavigna, M. (2012) *Discourse of Twitter and Social Media.* London: Continuum.

APPENDIX 4.1

COMPOSITION OF THE IMDB CORPUS

Thread Code	Message Board	Subject Line	Number of posts
001	Al Pacino	*Do you guys think he is a versatile actor?*	12
002	Audrey Hepburn	*Audrey or Marilyn? Who was more successful?*	16
003	Bruce Willis	*Assassin*	6
004	Bruce Willis	*It's time to retire, Brucie.*	20
005	Cary Grant	*Is it true he was bisexual?*	13
006	Clint Eastwood	*Dear Clint Eastwood;*	18
007	Henry Fonda	*Who looks like him?*	19
008	Henry Fonda	*Jane's anti War and America behavior*	8
009	Harrison Ford	*What happened to his career?*	15
010	Humphrey Bogart	*It's such a shame!*	16
011	James Stewart	*Mormon?*	2
012	James Stewart	*Did he have an affair?*	9
013	Kevin Spacey	*hes not gay is he?*	16
014	Kevin Spacey	*Midnight in the Garden of Good and Evil*	9
015	Marlon Brando	*Great Actor….It makes me smile….*	7
016	Marlon Brando	*Was he really bisexual?*	9
017	Robert De Niro	*Is It Time To Retire Already?*	22
018	Robert De Niro	*Robert's freakiest role????*	10
019	Takashi Shimura	*Great news for Shimura fans in the US*	6
020	Takashi Shimura	*What happend ? It is very strange!!*	4
021	Takashi Shimura	*Takashi Shimura*	6
022	Tom Cruise	*Tom Cruise makes Scientologists look like freaks!*	17
023	Johnny Depp	*I believe johnny depp is Santa Claus*	12
024	Johnny Depp	*could he ever match tom cruise?*	16
025	Angelina Jolie	*Am I the only teenage boy who finds her unattractive?*	18
026	Angelina Jolie	*her tattoos: Hot or Not?*	18
027	Brad Pitt	*Brad Pitt is aging real bad*	24
028	Vince Vaghn	*Is he really right-wing?*	20
029	George Clooney	*Confessions of a Dangerous Mind*	1
030	George Clooney	*Clooney + Bush Snr?*	15
031	Will Smith	*Will Smith refuses to become a scientologist*	12

032	Reese Witherspoon	*The first Withermovie i remember watching….*	23
033	Reese Witherspoon	*NOT WORTH 29 MILLION*	2
034	Adam Sandler	*Slowly getting funny and less funny and energetic??*	6
035	Adam Sandler	*I adore him/favorite Adam Sander movie?*	11
036	Adam Sandler	*Gay*	16
037	Tom Hanks	*Tom Hanks has to win another oscar!*	10
038	Tom Hanks	*In your Opinoin what is Tom Hanks Worst performance?*	9
039	Uma Thurman	*A guy that looks like UMA?*	12
040	Uma Thurman	*Last Uma Thurman film u watched???*	19
041	William Holden	*The truth about his death :(*	12
042	William Holden	*TO LORRAINE/EVERYBODY*	3
043	Russell Crowe	*Why is Russ's acting so bad in the AGY film?*	16
044	Jack Nicholson	*did he hate stanley kubrick?*	2
045	Jack Nicholson	*Big Time Castro Lover*	18
046	Joe Pesci	*Where has he been for the last 8 years?*	2
047	James Cameron	*Petition for Dark Angel movie*	18
048	Joel Coen	*The Coen Brothers problem….*	4
049	Joel Coen	*Arsenic and Old Lace*	1
050	Joel Coen	*Will they ever top Fargo?*	5
051	Joel Coen	*All of their film have been box-office flops*	5
052	Francis Ford Coppola	*How will he be remembered?*	4
053	Alfred Hitchcock	*Homophobia?*	14
054	Stanley Kubrick	*One of the, many, reasons I've come to not like Kubrick*	5
055	Stanley Kubrick	*Name Kubrick's weakest film*	17
056	Sergio Leone	*Did Sergio Leone ruin the Western?*	4
057	Gore Verbinski	*Is it me or is this guy all over the place with genres?*	11
058	Borat: Cultural Learnings of America for Make Benefit Glorious Nation of Kazakhstan (2006)	*If you like homo laden material....*	20
059	Alien (1979)	*This is the most overated sci fi film of all time!*	16
060	American Beauty (1999)	*Not to sound like a troll but I don't get…*	1
061	American Beauty (1999)	*Life changing*	11
062	American Beauty (1999)	*Ricky's father wasn't gay*	6
063	Band of Brothers (2001)	*Lugers*	10

064	Band of Brothers (2001)	*After World War II for Herbert Sobel*	6
065	Battlefield Earth: A Saga of the Year 3000 (2000)	*Am I the only one who liked that movie?*	26
066	Braveheart (1995)	*Why Braveheart Got a low Box Office?*	16
067	Braveheart (1995)	*The Worst 'Best' Oscar Movie Ever??*	6
068	Fight Club (1999)	*Why does Jack create Tyler?*	2
069	Fight Club (1999)	*really funny detail*	6
070	Fight Club (1999)	*Why is this rated so high?*	10
071	Forrest Gump (1994)	*Did Forrest and his kid have HIV?*	9
Total			***790***

APPENDIX 4.2
ANALYSIS OF ELEMENT RELATIONS IN THE *DE NIRO* THREAD

P	T	L	Element	Type
A	1	1.	Is It Time	|m–
		2.	to Retire	+m
		3.	already?	ms
		4.	C'mon	|oi
		5.	man,	|oi
		6.	if he keeps on making	|m–
		7.	mobster movies	+m
		8.	and	+ot–
		9.	bubble gum movies	ms
		10.	like	+ot–
		11.	Meet The Parents	ms
		12.	or	+ot–
		13.	Meet The Fockers,	ms
		14.	then	+ot–
		15.	that's just whack.	m
		16.	**Just tell me,**	**|oi**
		17.	**people,**	**|oi**
		18.	is it Time for him	|m–
		19.	to Retire?	+m
		20.	I wanna know.	m
B	2	21.	No!	m
C	3	22.	He has done	m–
		23.	some pretty crap films	+m
		24.	in the last 5 years.	ms
D	4	25.	De Niro will revive his career,	m
		26.	and	+ot–
		27.	**amke you guys**	**ms–**
		28.	**eat those words.**	**+m**
		29.	**To answer your question**	**|oi**
		30.	NO.	m
E	5	31.	like a boxer	|m
		32.	even great actors	m–
		33.	will work into their 80's	+m
		34.	and	+ot–

		35.	ridicule their legacy,	ms
		36.	brando-olivier	|m
		37.	they have all done it	m
F	6	38.	yeah	|oi
		39.	i couldnt agree with you more	m
		40.	buddy.	ois
		41.	this guy	m–
		42.	hasn't made a good film	+m
		43.	in along time.	ms
		44.	**i mean,**	**|oi**
		45.	he made those two films	m
		46.	godsend	m
		47.	and	+ot–
		48.	hide and seek.	ms
		49.	they were both	m–
		50.	the same goddamn movie.	+m
		51.	i think	|m
		52.	he should	m–
		53.	follow the example of	+m–
		54.	good actors	+m
		55.	like	+ot–
		56.	hugh grant	ms
		57.	and	+ot–
		58.	jude law	ms
		59.	and	+ot–
		60.	choose his roles	ms–
		61.	a bit better	+m
G[124]	7	62.	[quote] i think	|m
		63.	he should	m–
		64.	follow the example of	+m–
		65.	good actors	+m
		66.	like	+ot–
		67.	hugh grant	ms
		68.	and	+ot–

[124] Shaded elements are those which are quoted using the quotation markup.

		69.	jude law	ms
		70.	and	+ot–
		71.	choose his roles	ms–
		72.	a bit better [/quote]	+m
		73.	**ROTFLMAO!!!**	**oi**
H	8	74.	hey,	|oi
		75.	hugh grant and jude law	m–
		76.	are great actors!	+m
		77.	they made some fantastic movies	m
		78.	like	+ot–
		79.	euh...	|oi
		80.	hmm...	|oi
		81.	ok	|oi
		82.	**wait...**	**oi**
		83.	hmm...	|oi
		84.	**no wait**	**|oi**
		85.	i got one...	oi
		86.	hmm	|oi
		87.	no....	oi
		88.	**ROFL**	**oi**
I	9	89.	"even great actors	m–
		90.	will work into their 80's	+m
		91.	and	+ot–
		92.	ridicule their legacy,"	ms
		93.	**First off**	**+ot–**
		94.	their legacy is not tarnished	m
		95.	because	+ot–
		96.	once they die	|m
		97.	the bad films	m–
		98.	will be forgotten	+m
		99.	anyway.	ois
		100.	**Second,**	**+ot–**
		101.	we need someone	m–

		102.	to play the parts	+m
		103.	for older people...	ms
		104.	or	+ot–
		105.	do they not even exist	|m
		106.	to you?	ms
E	10	107.	of course	|oi
		108.	they ruin their legacy,	m
		109.	when they ham up	|m
		110.	in rubbish films	ms
		111.	they make themselves	m–
		112.	look like fools,	+m
		113.	when Olivier did Inchon-Jazz Singer	|m
		114.	etc	ois
		115.	he was absolutely	m–
		116.	making a fool of himself	+m
		117.	and	+ot–
		118.	looked ridiculous.	ms
		119.	so	+ot–
		120.	**to say**	**|m**
		121.	you dont ruin your legacy	m–
		122.	is rubbish	+m
		123.	because	+ot–
		124.	you do.	m
		125.	Pacino selects	m–
		126.	what he does	+m
		127.	v carefully	ms
	11	128.	<post deleted>	
J	12	129.	"Gigli	m
		130.	and	+ot–
		131.	Simone	ms
		132.	anyone?"	ois
		133.	Insomnia,	m

		134.	People I Know,	ms
		135.	The Insider,	ms
		136.	Angels in America,	ms
		137.	and	+ot–
		138.	The Merchant of Venice	ms
		139.	anyone?	ois
E	13	140.	[quote] Insomnia,	m
		141.	People I Know,	ms
		142.	The Insider,	ms
		143.	Angels in America,	ms
		144.	and	+ot–
		145.	The Merchant of Venice	ms
		146.	anyone? [/quote]	ois
		147.	they are all good movies,	m
		148.	but	+ot–
		149.	perhaps	|oi
		150.	you should	m–
		151.	go watch Hide And Seek	+m
		152.	again	ms
		153.	eh	ois
J	14	**154.**	**Let's take a look at**	**m–**
		155.	what Pacino's done	+m
		156.	in the last 10 years:	ms
		157.	Donnie Brasco,	m
		158.	The Devil's Advocate,	ms
		159.	The Insider,	ms
		160.	Insomnia,	ms
		161.	Angels in America,	ms
		162.	People I Know,	ms
		163.	The Merchant of Venice,	ms
		164.	etc.	ois
		165.	All great movies	m

		166.	with great performances from Pacino.	ms
		167.	**Now**	**|oi**
		168.	**let's take a look at**	**m–**
		169.	what De Niro's done	+m
		170.	in the past 10 years:	ms
		171.	Cop Land,	m
		172.	The Adventures of Rocky & Bullwinkle,	ms
		173.	15 Minutes,	ms
		174.	The Score,	ms
		175.	Showtime,	ms
		176.	City by the Sea,	ms
		177.	Analyze That,	ms
		178.	Godsend,	ms
		179.	The Bridge of San Luis Rey,	ms
		180.	etc.	ois
		181.	Mostly bad to average movies	m
		182.	with equally bad	ms–
		183.	to average performances.	+m
		184.	**Nothing special here.**	**m**
		185.	The main difference between Pacino	|m–
		186.	and	+ot–
		187.	De Niro	+m–
		188.	lately	|m
		189.	is	+m–
		190.	that	+ot–
		191.	Pacino's last great performance	m–
		192.	was 2 years ago,	+m
		193.	in The Merchant of Venice,	ms
		194.	and	+ot–
		195.	De Niro's last great performance	m–
		196.	was 11 years ago,	+m
		197.	in Heat.	ms
A	15	198.	I agree big time	m

		199.	with pianoman17.	ms
		200.	**LOL**	**|oi**
		201.	The most recent movies	m–
		202.	that DeNiro have been making	|m
		203.	had been a joke!	+m
		204.	And	+ot–
		205.	the most recent movies	m–
		206.	that Pacino made	|m
		207.	were actually	+m–
		208.	carefully selected	+m
		209.	(impressed).	ms
		210.	Not bad at all for Pacino.	m
	16	211.	<post deleted>	
J	17	212.	Really?	|oi
		213.	'Cause	+ot–
		214.	I remember very clearly	|m
		215.	that	+ot–
		216.	you've included Pacino	m–
		217.	in your "best actors" list	+m
		218.	on several occasions.	ms
	18	219.	<post deleted>	
J	19	220.	You can think otherwise	m
		221.	all you want,	ms
		222.	but in reality	+ot–
		223.	you did.	m
K	20	224.	Jack Nicholson has done	m–
		225.	a better job	+m
		226.	for the past 10 years	ms
		227.	than	+ot–
		228.	Pacino	ms
		229.	or	+ot–
		230.	DeNiro.	ms

L	21	231.	He's obviously	m–
		232.	lost it	+m
		233.	He cant pull off	m–
		234.	the gangster/hardman roles	+m
		235.	anymore	ms
		236.	that hes been	ms–
		237.	so good at	+m
		238.	so	+ot–
		239.	hes tryin his hand	m
		240.	at comedy	ms
		241.	even though	+ot–
		242.	he doesnt have	m–
		243.	a funny bone in his body.	+m
A	22	244.	Yeah,	loi
		245.	I agre with you	m
		246.	man.	ois
		247.	He's too old to play	m–
		248.	those gangster/hardman roles.	+m
		249.	He sucks now,	m
		250.	man,	ois
		251.	he needs to either retire	m–
		252.	or	+ot–
		253.	need to	+m–
		254.	pick roles his age,	+m
		255.	dammit.	ois
M	23	256.	Deniro still the man!	m
		257.	he can	m–
		258.	pull out any roll	+m
		259.	at his age	ms
N	24	260.	Deniro is brilliant..	m
		261.	his acting was great	m–
		262.	in Hide and Seek..	+m

		263.	very believable.	m
		264.	Deniro can play	m–
		265.	a diverse range of roles	+m
		266.	and	+ot–
		267.	gives any movie	ms–
		268.	an immediate sense of grandeur..	+m
		269.	in much the same way as	+ot–
		270.	Al Pacino..	ms
		271.	keep up the awesome work!!	m
O	25	272.	what does deniro	m–
		273.	have to retire	+m
		274.	if he likes doing	ms–
		275.	what he does:	+m
		276.	acting.	m
		277.	So what if	loi
		278.	his career lately	m–
		279.	hasn't been	+m–
		280.	as great as before,	+m
		281.	but	+ot–
		282.	it happens	m
		283.	with every actor	ms
		284.	in their career.	ms
		285.	As of now,	+ot–
		286.	maybe	loi
		287.	he wants to	m–
		288.	try different roles	+m
		289.	and	+ot–
		290.	experiment,	ms
		291.	he's been around the buisness	m
		292.	for quite awhile	ms
		293.	and	+ot–
		294.	has done various roles.	ms
		295.	However,	+ot–

		296.	he shouldn't retire,	m
		297.	because	+ot–
		298.	he enjoys acting	m
		299.	and	+ot–
		300.	if that's	|m–
		301.	what makes him happy	+m
		302.	why not	m–
		303.	continue doing it?	+m
P	26	304.	NO !!!!!	|oi
		305.	that's Robert De Niro.	m
		306.	not	ms–
		307.	one of these B-class actors.	+m
		308.	he was,	m–
		309.	is	|m–
		310.	and	+ot–
		311.	will ever be	+m–
		312.	one of the greatest actors	+m
		313.	in Film History.	ms
		314.	Hands down!	oi
Q	27	315.	De Niro is	m–
		316.	the most admired actor	+m
		317.	currently working today.	ms
		318.	FACT.	oi
		319.	No chance of him retiring.	m

APPENDIX 4.3
ANALYSIS OF LINEAR UNIT RELATIONS IN THE *DE NIRO* THREAD

P	T	L	Linear Unit	Type	Retrospective supplementary mechanism	Structure (elements)
A	1	1.	Is It Time / To Retire / Already?	M–		|m– / +m / ms
		2.	C'mon / man, if he keeps on making / mobster movies / and / bubble gum movies / like / Meet The Parents / or / Meet The Fockers,/ then / that's just whack.	|M	Encap. (implicit)	|oi /|oi |m– / +m / +ot– / ms / +ot– / ms / +ot– / ms/ +ot– / m
		3.	**Just tell me, / people,** is it Time for him / to Retire?	M–		**|oi / |oi** / |m– / +m
		4.	I wanna know.	|OI	Encap. (implicit)	m
B	2	5.	NO!	+M		m
C	3	6.	He has done / some pretty crap films / in the last 5 years.	+M		m– / +m / ms
D	4	7.	De Niro will revive his career, / and / amke you guys / **eat those words.**	**|M**	Encap. (explicit)	m / +ot– / **ms– / +m**
		8.	**To answer your question /** NO	**MS**	**Overlay**	**|oi /** m
E	5	9.	like a boxer / even great actors / will work into their 80's / and / ridicule their legacy,	+M		|m / m– / +m / +ot– / ms
		10.	brando-olivier / they have all done it	MS	Encap. (fuzzy)	|m / m
F	6	11.	yeah / i couldnt agree with you more / buddy.	MS–	Encap. (implicit)	|oi / m /ois
		12.	this guy / hasn't made a good film / in along time.	+M		m– / +m / ms
		13.	**i mean,/** he made those two films	MS–	**Encap. (implicit)**	**|oi** m
		14.	godsend / and / hide and seek.	+M–		m / +ot– / ms
		15.	they were both / the same goddamn movie.	+M		m– / +m
		16.	i think / he should / follow the example of / good actors / like / hugh grant / and / jude law/ and / choose his roles / a	M		|m / m– / +m– / +m/ +ot– / ms/ +ot– / ms / +ot– /

			bit better			ms– / +m
G	7	17.	[quote] i think he should follow the example of good actors like hugh grant and jude law and choose his roles a bit better [/quote]	|OI	Verbal echo	|m / m– / +m– / +m/ +ot– / ms/ +ot– / ms / +ot– / ms– / +m
		18.	**ROTFLMAO!!!**	**|OI**	**Encap. (implicit)**	**oi**
H	8	19.	hey, /hugh grant and jude law / are great actors!	|M–	Overlay	|oi / m– / +m
		20.	they made some fantastic movies / like	+M–		m / +ot–
		21.	euh…/ hmm… / ok / **wait…**	|OI	Encap. (implicit)	|oi / |oi / |oi / **oi**
		22.	hmm... / **no wait** / i got one...	**|OI**	Encap. (implicit)	|oi / **|oi / oi**
		23.	hmm / no....	|OI	Encap. (implicit)	|oi / oi
		24.	ROFL	OIS	Encap. (implicit)	oi
I	9	25.	"even great actors will work into their 80's and ridicule their legacy"	|OI	Verbal echo	/ m– / +m / +ot– / ms
		26.	**First off /** their legacy is not tarnished	**|M–**	**Overlay**	**+ot–/** m
		27.	because /once they die / the bad films / will be forgotten / anyway.	+M		+ot–/ |m / m– / +m /ois
		28.	**Second, /** we need someone / to play the parts / for older people...	**MS–**	**Encap. (implicit)**	**+ot– /**m– /+m / ms
		29.	or /do they not even exist / to you?	+M–		+ot– / |m / ms
E	10	30.	of course / they ruin their legacy,	|M–	Overlay	|oi / m
		31.	when they ham up / in rubbish films / they make themselves / look like fools,	+M		|m / ms / m– / +m

		32.	when Olivier did Inchon-Jazz Singer / etc / he was absolutely / making a fool of himself / and / looked ridiculous.	MS	Overlay	|m / ois / m– / +m / +ot– / ms
		33.	so **/ to say** / you dont ruin your legacy / is rubbish	**MS**	**Overlay**	+ot– **/ |m /** m - / +m
		34.	because / you do.	MS	Encap. (implicit)	+ot– / m
		35.	Pacino selects / what he does / v carefully	M		m– / +m / ms
	11	36.	post deleted			
J	12	37.	"Gigli and Simone anyone?"	|OI	Verbal echo	m/ +ot– / ms / ois
		38.	Insomnia, / People I Know, / The Insider, / Angels in America, / and / The Merchant of Venice / anyone?	|M–	Verbal echo	m / ms / ms / ms / +ot– / ms / ois
E	13	39.	[quote] Insomnia, People I Know, The Insider, Angels in America, and The Merchant of Venice anyone? [/quote]	|OI	Verbal echo	m/ ms /ms/ ms /+ot–/ ms/ ois
		40.	they are all good movies,	+M–	Encap. (fuzzy)	m
		41.	but / perhaps / you should / go watch Hide And Seek / again/ eh	+M–	Encap. (implicit)	+ot–/ |oi / m– / +m / ms / ois
J	14	42.	**Let's take a look at /** what Pacino's done / in the last 10 years:	**M -**		**m–** / +m / ms
		43.	Donnie Brasco, / The Devil's Advocate, / The Insider, / Insomnia, / Angels in America, / People I Know, / The Merchant of Venice, / etc.	+M		m / ms / ms / ms / ms / ms / ms / ms / ois
		44.	All great movies / with great performances from Pacino.	MS	Encap. (implicit)	m / ms

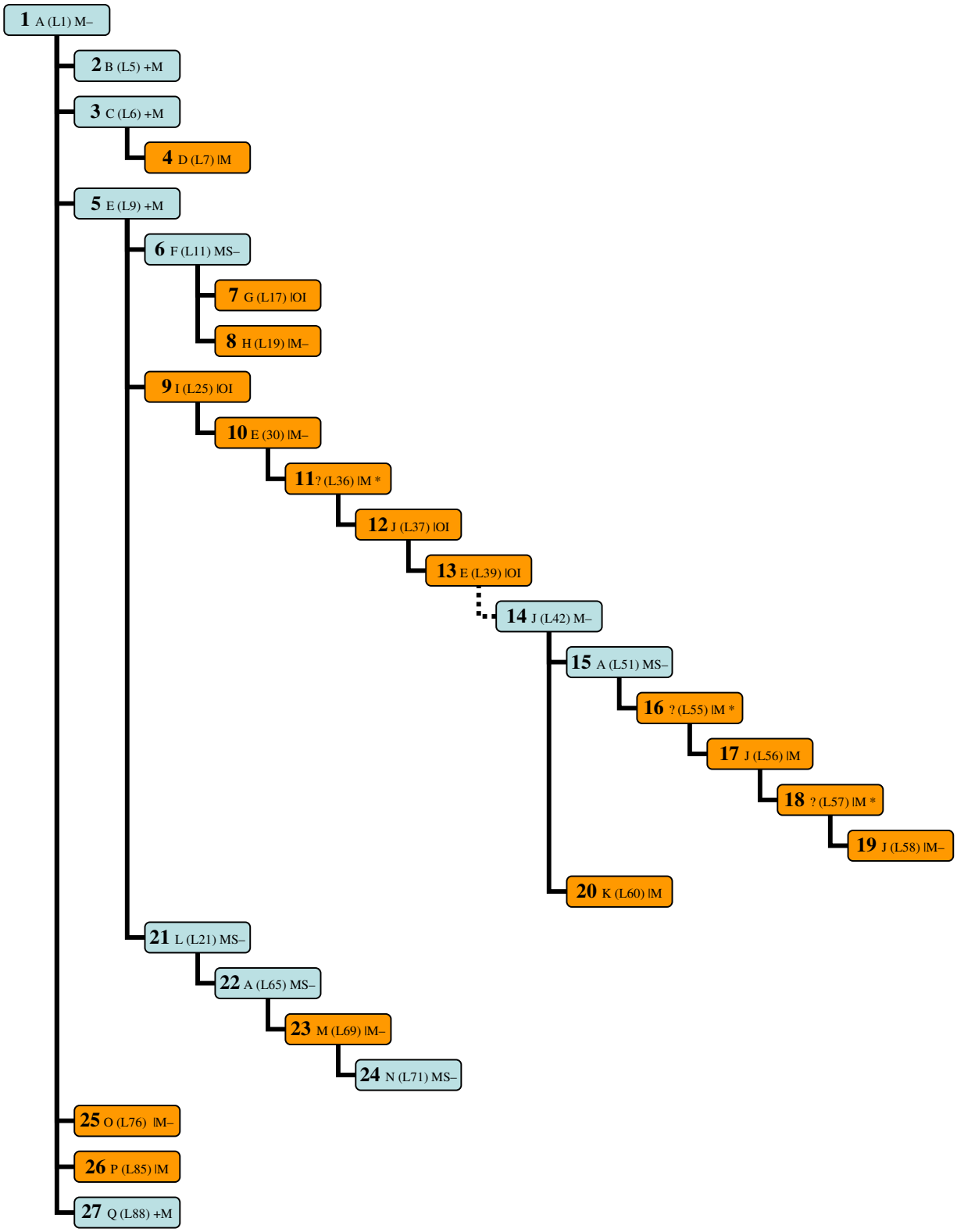
		45.	**Now / Let's take a look at /** what De Niro's done / in the past 10 years:	**MS–**	Encap. (implicit)	**|oi / m–** / +m / ms
		46.	Cop Land, / The Adventures of Rocky & Bullwinkle, / 15 Minutes, / The Score, / Showtime, / City by the Sea, / Analyze That, / Godsend, The Bridge of San Luis Rey, etc	+M		m / ms / ms / ms /ms / ms /ms /ms / ms / ois
		47.	Mostly bad to average movies / with / equally bad to average performances.	MS	Encap. (implicit)	m / ms– / +m
		48.	**Nothing special here.**	**MS**	**Overlay**	**m**
		49.	The main difference between Pacino / and / De Niro / lately / is /that / Pacino's last great performance / was 2 years ago, / in The Merchant of Venice,	M–		|m– / +ot– /+m– / |m / +m– / +ot– / m– / +m / ms
		50.	and / De Niro's last great performance / was 11 years ago, / in Heat.	+M		+ot– / m– / +m / ms
A	15	51.	I agree big time / with <name>.	MS–	Encap. (implicit)	m /ms
		52.	**LOL** / The most recent movies / that DeNiro have been making / had been a joke!	+M		**|oi** / m– / |m /+m
		53.	And / the most recent movies / that Pacino made / were actually / carefully selected / (impressed).	MS	Overlay	+ot– / m– / |m / +m– / +m /ms
		54.	Not bad at all for Pacino.	MS	Overlay	m
	16	55.	post deleted			
J	17	56.	Really? 'Cause / I remember very clearly / that / you've included Pacino / in your "best actors" list / on several	|M	Encap. (implicit)	|oi / +ot– / m / +ot– / m– / +m / ms

			occasions.			
	18	57.	post deleted			
J	19	58.	You can think otherwise / all you want,	|M–	Encap. (implicit)	m / ms
		59.	but in reality / you did.	-M		+ot– /m
K	20	60.	Jack Nicholson has done / a better job / for the past 10 years than Pacino / or / DeNiro.	|M	Overlay	m - / +m / ms / +ot– / ms / +ot– / ms
L	21	61.	He's obviously / lost it.	MS–	Overlay	m– / +m
		62.	He cant pull off / the gangster/hardman roles / anymore / that hes been / so good at	+M		m– / +m / ms / ms– / +m
		63.	so / hes tryin his hand / at comedy	MS	Encap. (implicit)	+ot– / m / ms
		64.	even though / he doesnt have / a funny bone in his body.	MS	Encap. (implicit)	+ot– / m– / +m
A	22	65.	Yeah, / I agre with you / man.	MS–	Encap. (implicit)	|oi / m / ois
		66.	He's too old to play / those gangster/hardman roles.	+M		m– / +m
		67.	He sucks now, / man,	MS–	Overlay	m / ois
		68.	he needs to either retire / or / need to / pick roles his age, / dammit.	+M		m– / +ot– / +m–/ +m / ois
M	23	69.	Deniro still the man!	|M–	Overlay	m
		70.	he can / pull out any roll / at his age	+M		m– / +m / ms
N	24	71.	Deniro is brilliant..	MS–	Overlay	m
		72.	his acting was great / in Hide and Seek..	+M		m– / +m
		73.	very believable.	MS	Encap. (implicit)	m

		74.	Deniro can play / a diverse range of roles / and / gives any movie / an immediate sense of grandeur../ in much the same way as / Al Pacino..	M		m– / +m / +ot– / ms– / +m / +ot– / ms
		75.	keep up the awesome work!!	MS	Encap. (explicit)	m
O	25	76.	what does deniro / have to retire / if he likes doing / what he does:	|M–	Overlay	m– / +m / ms– / +m
		77.	acting.	+M		m
		78.	So what if/ his career lately / hasn't been / as great as before,	M		|oi / m– / +m– / +m
		79.	but / it happens / with every actor / in their career.	MS	Encap. (fuzzy)	+ot– /m / ms / ms
		80.	As of now / maybe / he wants to / try different roles / and / experiment,	MS–	Encap. (implicit)	+ot– / |oi/ m– / +m / +ot– / ms
		81.	he's been around the buisness / for quite awhile / and / has done various roles.	+M		m / ms / +ot– / ms
		82.	However, he shouldn't retire,	MS–	Encap. (implicit)	+ot– /m
		83.	because he enjoys acting	+M		+ot– / m
		84.	and / if that's / what makes him happy / why not / continue doing it?	MS	Encap. (fuzzy)	+ot– / |m– / +m / m– / +m
P	26	85.	NO !!!!! /that's Robert De Niro. / not / one of these B-class actors.	|M	Encap. (fuzzy)	|oi / m / ms– / +m
		86.	he was, / is / and / will ever be / one of the greatest actors / in Film History.	MS	Overlay	m - / |m– / +ot– / +m– / +m / ms

		87.	Hands down!	OIS	Encap. (implicit)	oi
Q	27	88.	De Niro is / the most admired actor / currently working today.	+M		m– / +m / ms
		89.	FACT.	OIS	Encap. (implicit)	oi
		90.	No chance of him retiring.	M		m

APPENDIX 4.4: THE *DE NIRO* THREAD NEST STRUCTURE



* = Deleted post. The linear compliance is guessed from the subsequent turn

Solid line between shapes means linear structural connection. Dotted line means no linear structural connection but there is a notional connection.

APPENDIX 4.5

INSTRUCTIONS FOR ANNOTATION SCHEMES IN THE *UAM CORPUSTOOL*

A4.5.1 Annotation scheme for element relations

Step 1: Creating Layer

i. Add an annotation layer to project called *Element.*

ii. *Coding Object*: Choose: *Annotate segments*.

iii. *What kind of segments are you creating?* Choose: *Plain text segments.*

iv. *Coding Scheme*: Choose: *Create new scheme.*

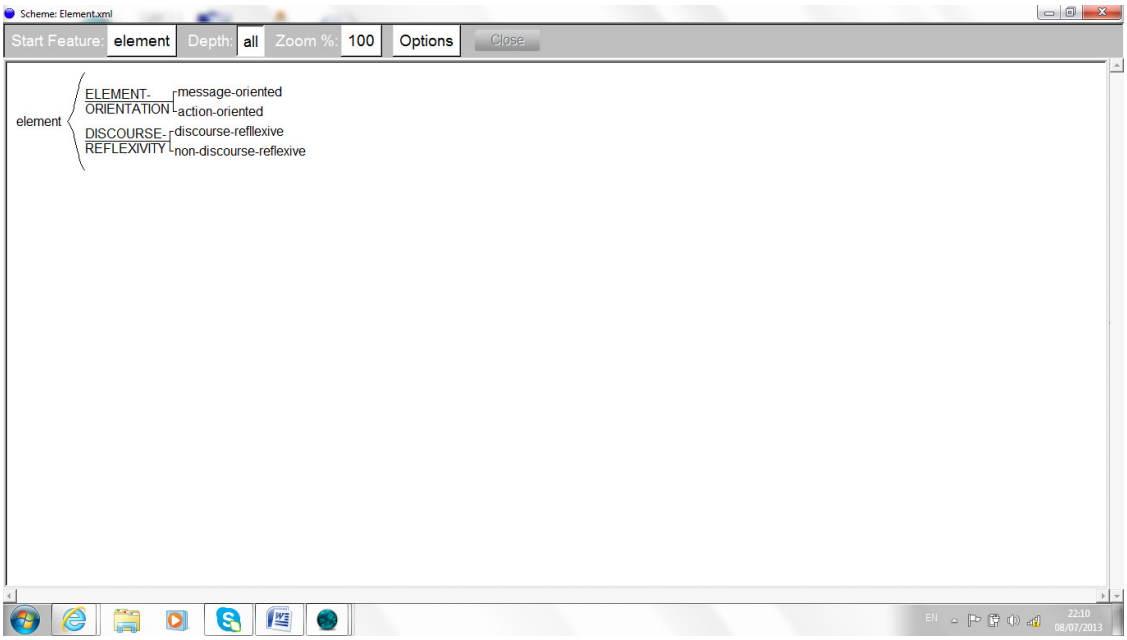
Step 2: Creating the two systems: element orientation and discourse reflexivity

i. Rename system: *ELEMENT ORIENTATION.*

ii. Add system: *DISCOURSE REFLEXIVITY.*

iii. Rename features: *message oriented* and *action oriented.*

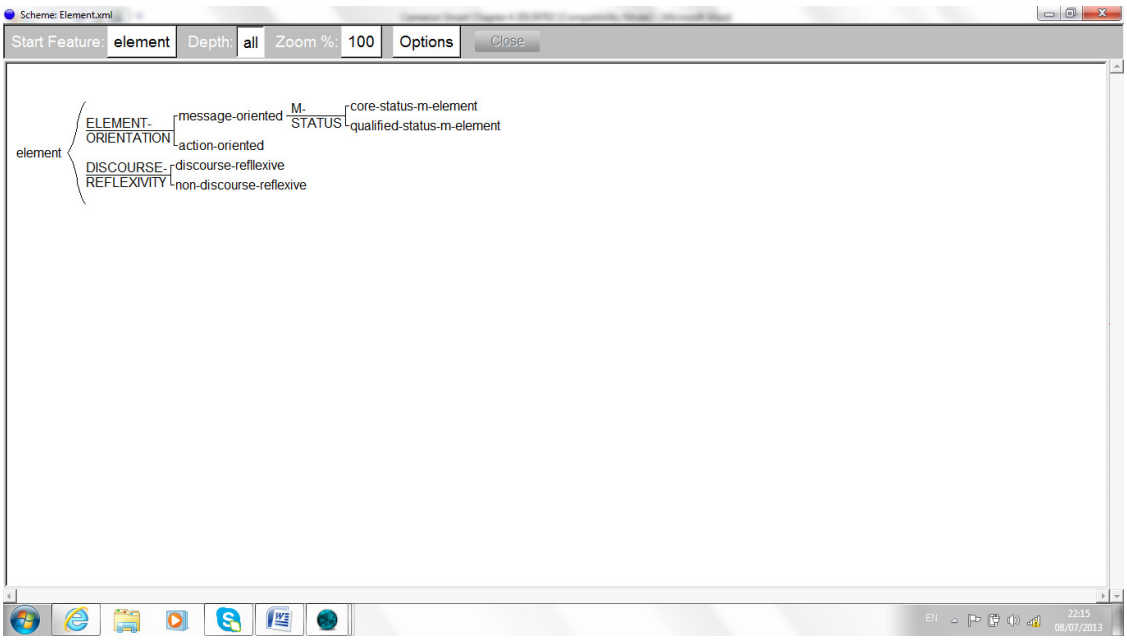
iv. Rename features: *discourse reflexive* and *non discourse reflexive.*



Step 3: Creating core and qualified m elements

i. Add system to *message oriented*. Rename it : *M STATUS.*

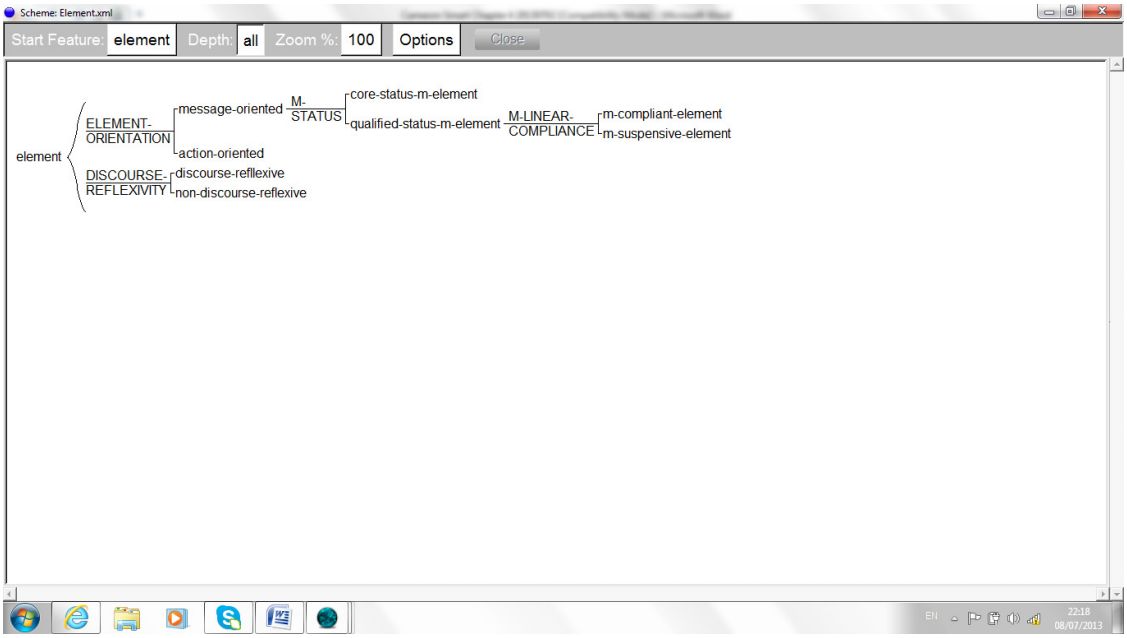
ii. Rename features: *core status m element* and *qualified status m element.*



Step 4: Creating compliant and suspensive m elements

i. Add system to *qualified status m element*. Rename the system *M LINEAR COMPLIANCE.*

ii. Rename features: *compliant m element* and *suspensive m element.*



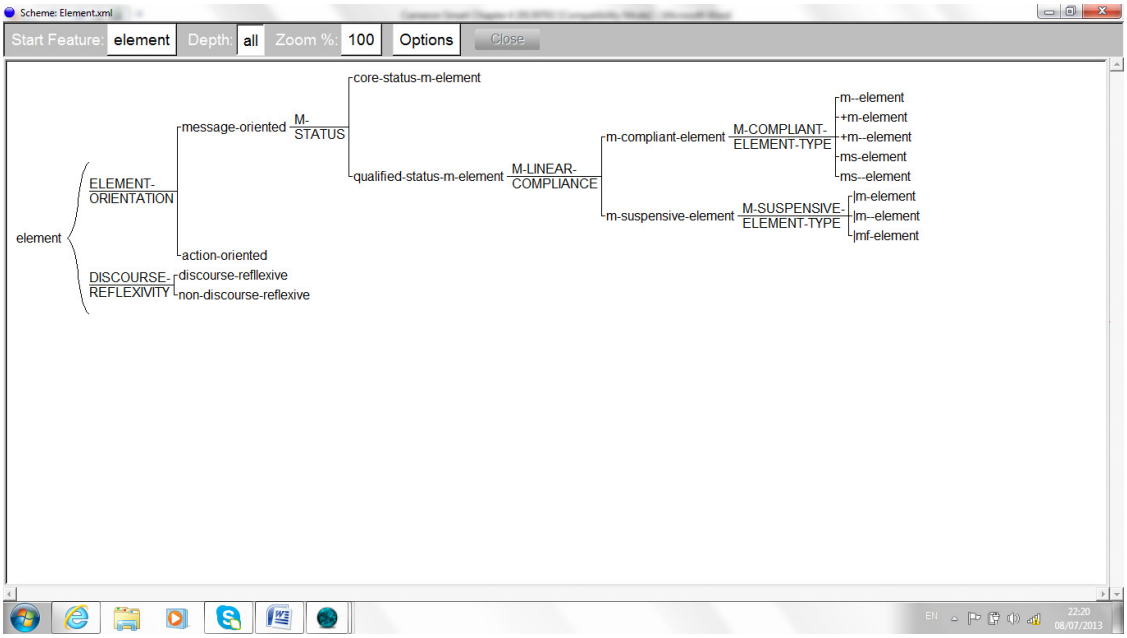
Step 5: Adding subcategories of m elements

i. Add system to *compliant m element.*

ii. Rename and add features: *m– element*, *+m element*, *+m– element*, *ms element*, *ms– element.*

iii. Add system to *suspensive m element.*

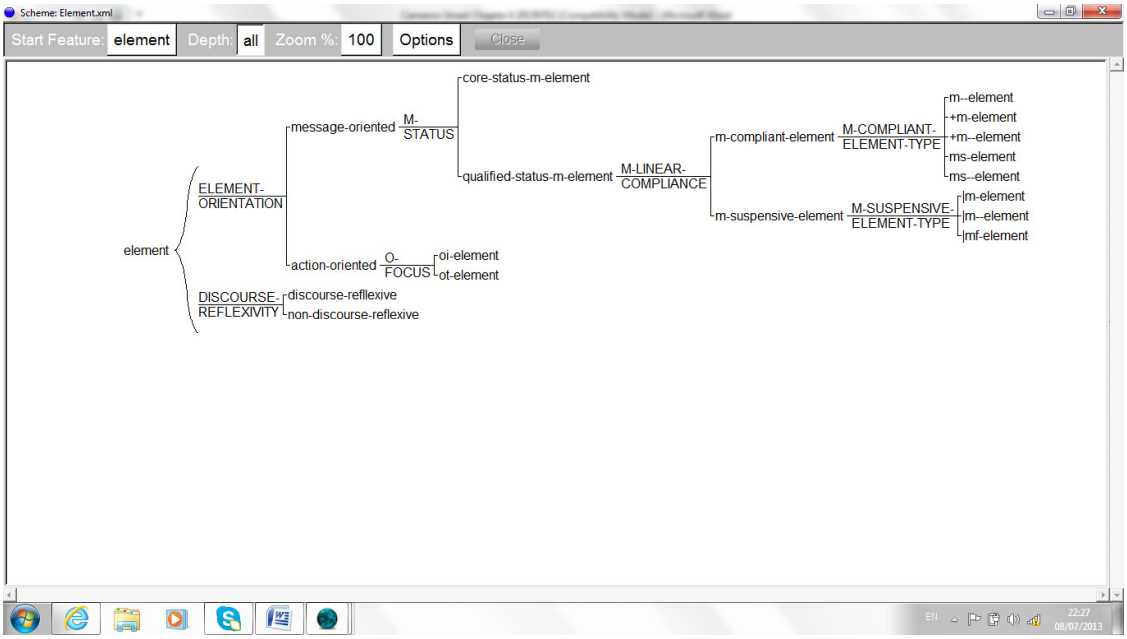
iv. Rename and add features: *|m element*, *|m– element*, *|mf element.*



Step 6: Creating oi and ot elements

i. Add system to *action oriented*. Rename: *O FOCUS.*

ii. Rename features: *oi element* and *ot element.*

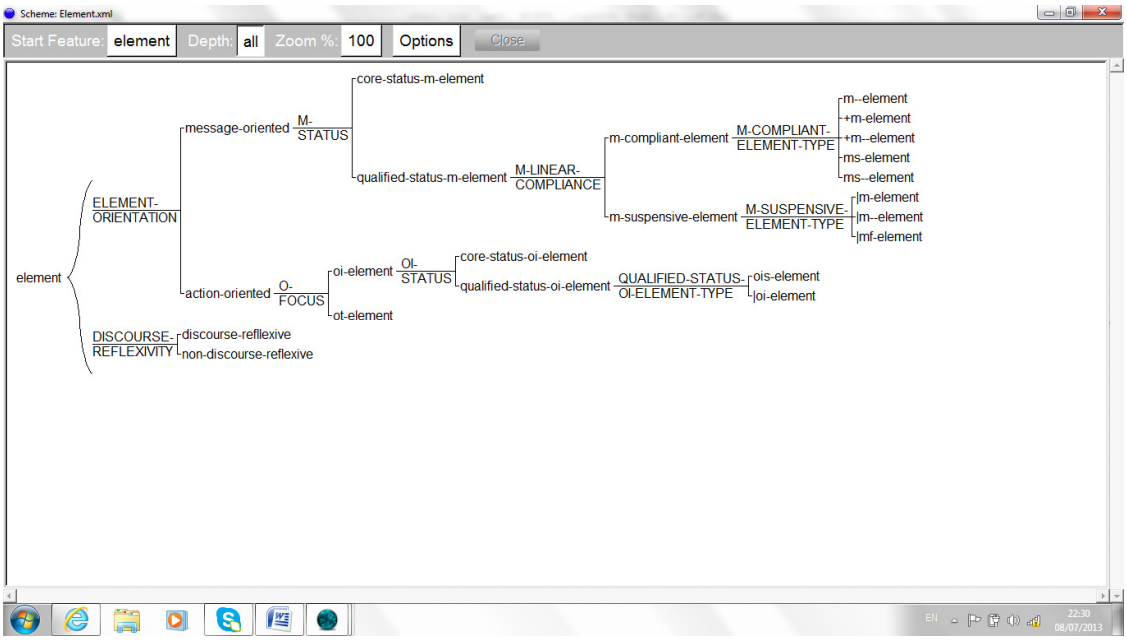


Step 7: Creating subcategories of oi elements

i. Add system to oi element. Rename system: OI STATUS.

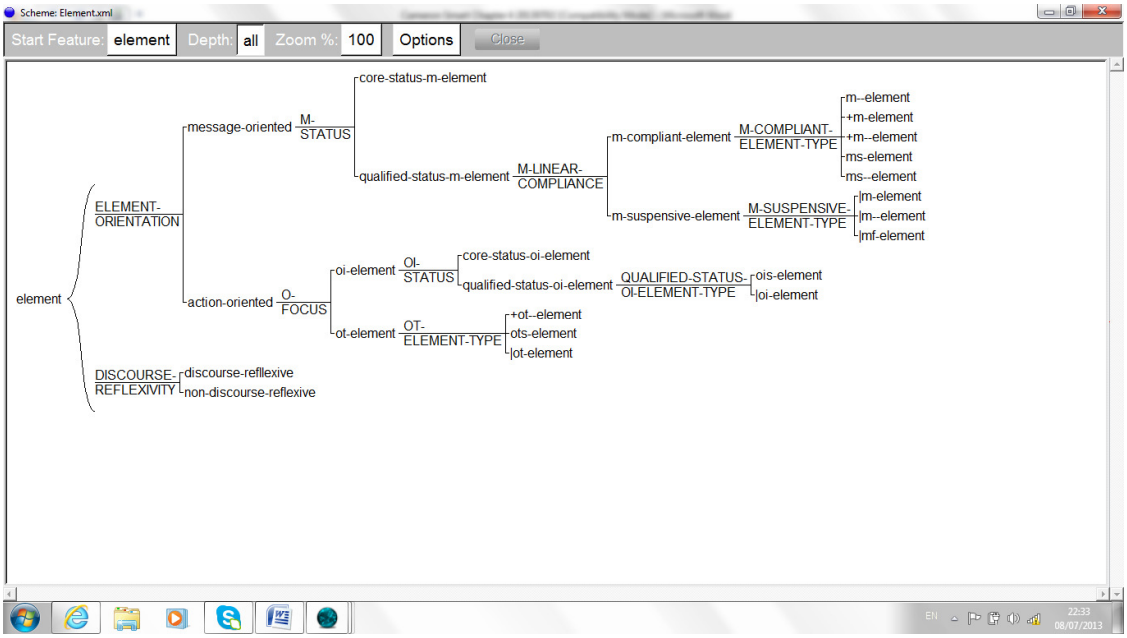
ii. Rename features: *core status oi element* and *qualified status oi element.*

iii. Add system to *qualified status oi element*. Rename features: *ois element* and *|oi element.*



Step 8: Creating subcategories of ot elements

i. Add system to *ot element*. Rename and add features: *+ot– element*, *ots element*, *|ot element.*



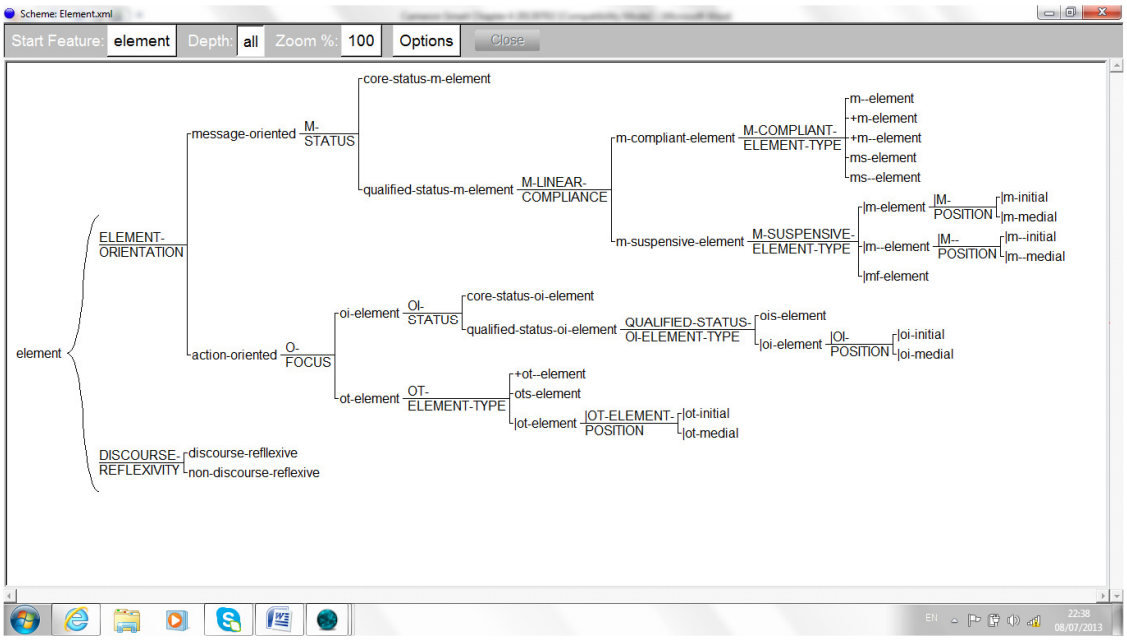
Step 9: Creating position in linear unit information for suspensive elements

i. Add systems to |*m element*. Rename the features: |*m initial* and |*m medial.*

ii. Add systems to |*m– element*. Rename the features: |*m– initial* and |*m– medial.*

iii. Add systems to |*oi element*. Rename the features: |*oi initial* and |*oi medial.*

iv. Add systems to |*ot element*. Rename the features: |*ot initial* and |*ot medial.*



A4.5.2 Annotation scheme for linear unit relations

Step 1: Creating Layer

i. Add an annotation layer to project called *Linear unit.*

ii. *Coding Object*: Choose: *Annotate segments*.

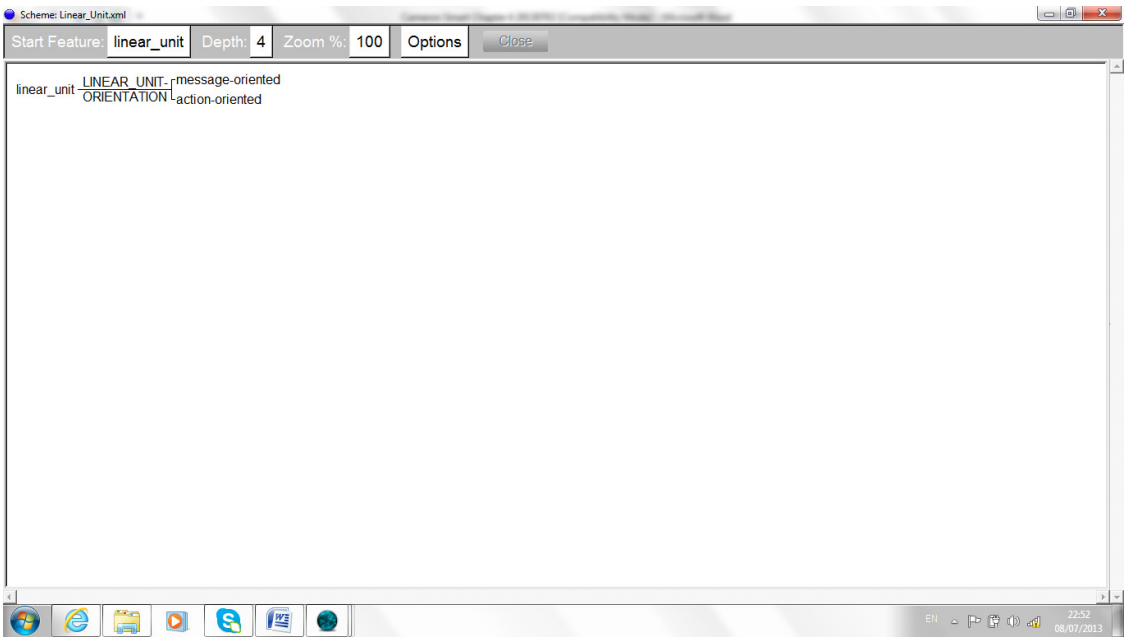
iii. *What kind of segments are you creating?* Choose: *Plain text segments.*

iv. *Coding Scheme*: Choose: *Create new scheme.*

Step 2: Creating the two linear unit categories

i. Rename the system LINEAR UNIT ORIENTATION.

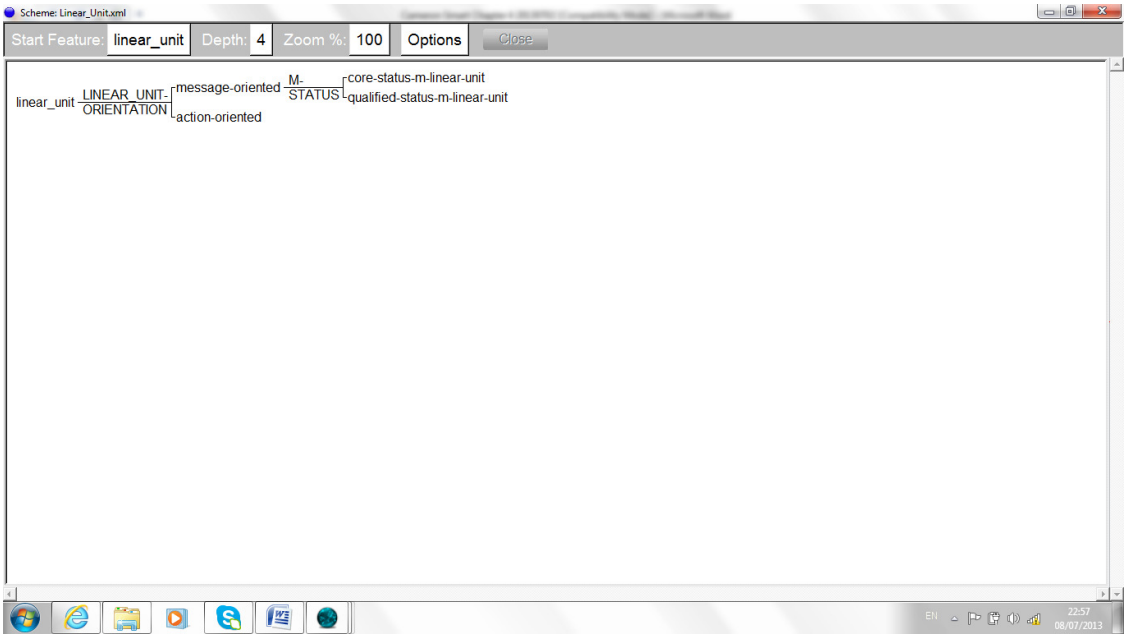
ii. Rename the features: *message oriented* and *action oriented.*



Step 3: Creating core and qualified M linear units

i. Add system to *message oriented*. Rename the system M STATUS.

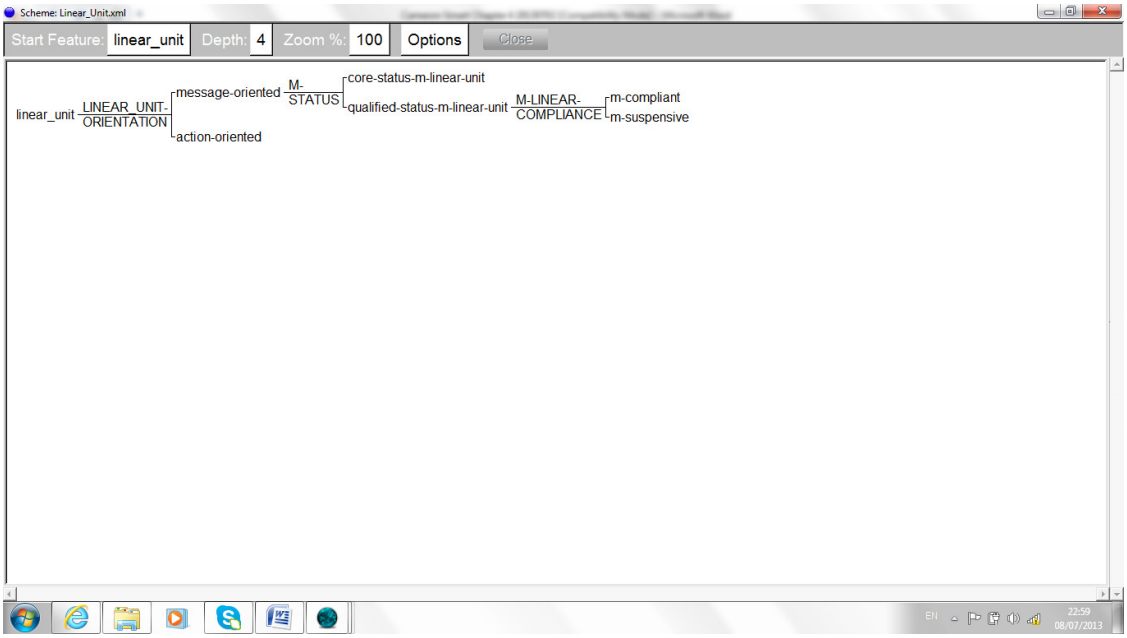
ii. Rename the features: *core m linear unit* and *qualified status m linear unit*.



Step 3: Creating compliant and suspensive M linear units

i. Add system to qualified status m linear unit. Rename the system: M LINEAR *COMPLIANCE.*

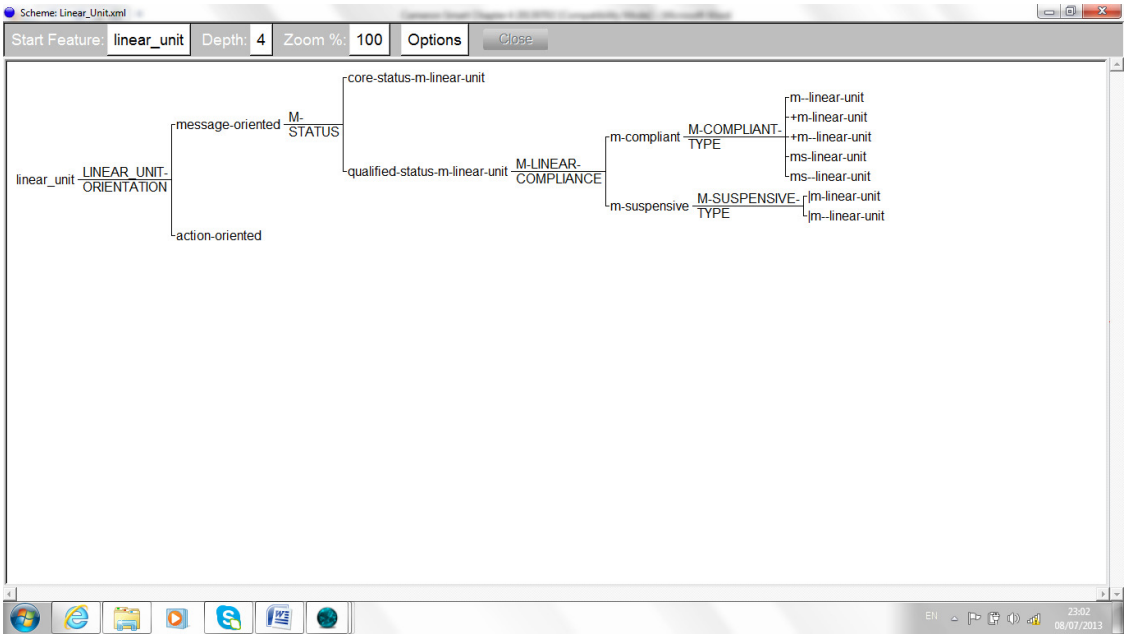
ii. Rename features: *m compliant* and *m suspensive.*



Step 4: Creating subcategories of M linear units

i. Add system to *m compliant*. Rename and add features: *m– linear unit*, *+m linear unit*, *+m– linear unit*, *ms linear unit* and *ms– linear unit*.

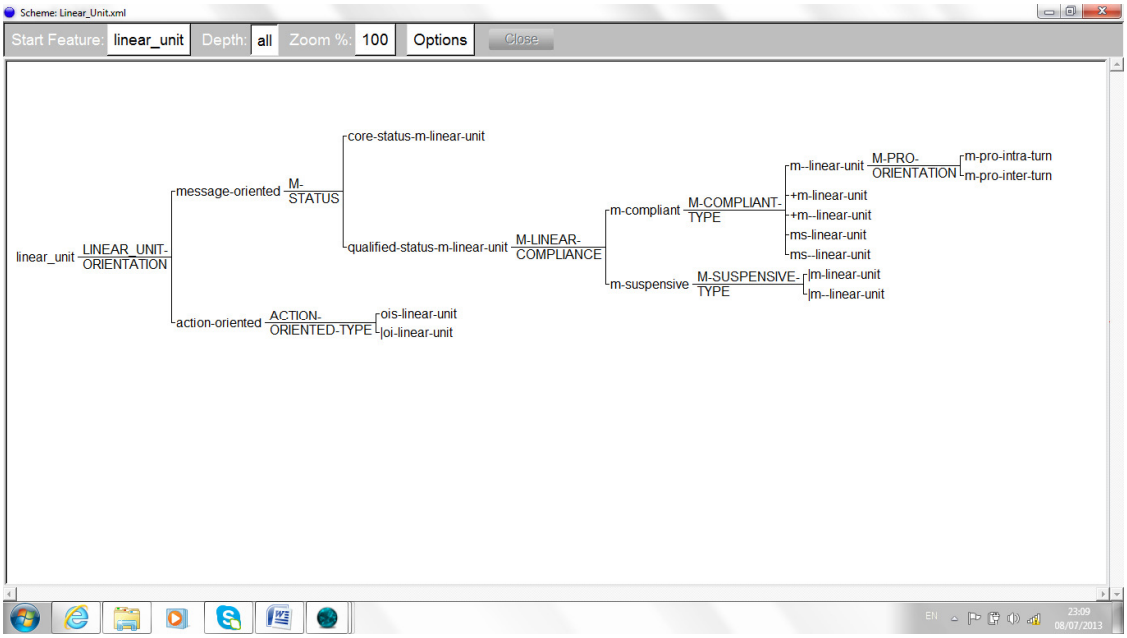
ii. Add system to *m suspensive*. Rename features: *|m linear unit* and *|m– linear unit.*



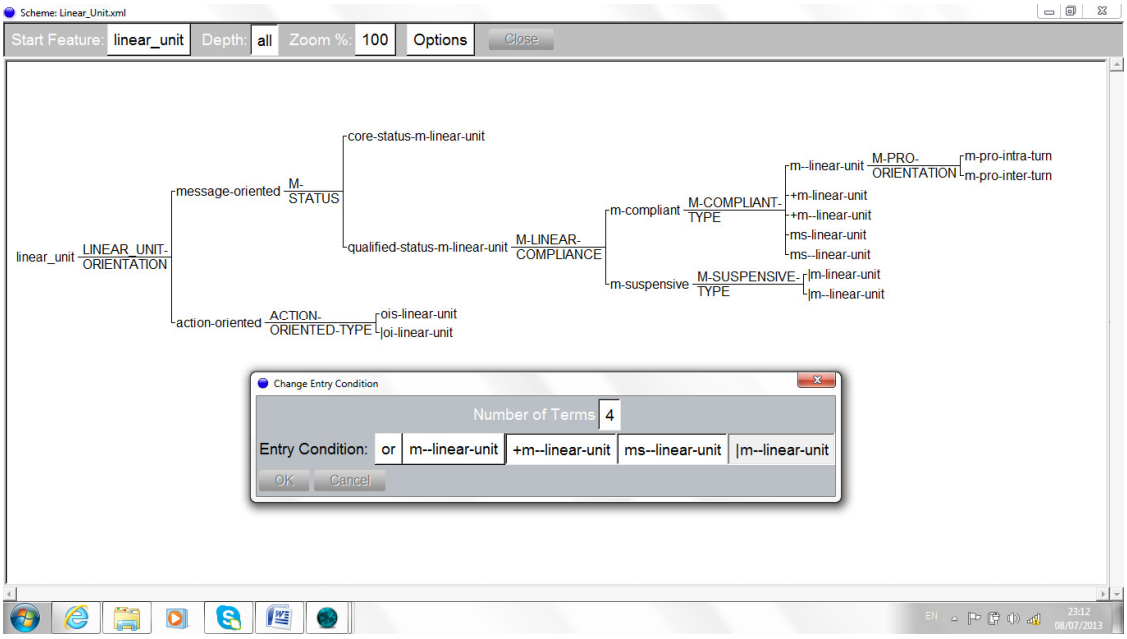
Step 5: Creating pro turn orientation for M linear units

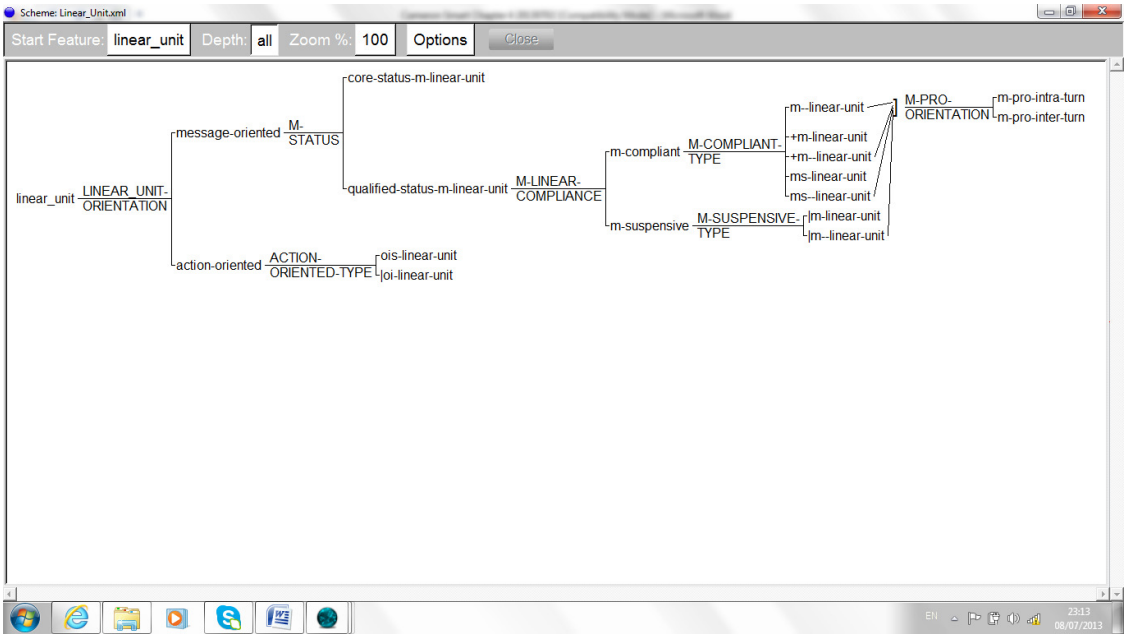
i. Add system to *m– linear unit*. Rename the system *M PRO ORIENTATION.*

ii. Rename the features: *m pro intra turn* and *m pro inter turn.*



iii. Change Entry Condition of M PRO ORIENTATION system: Number of terms: 4; Entry condition: *or m– linear unit*, *+m– linear unit*, *ms– linear unit*, *|m– linear unit*.



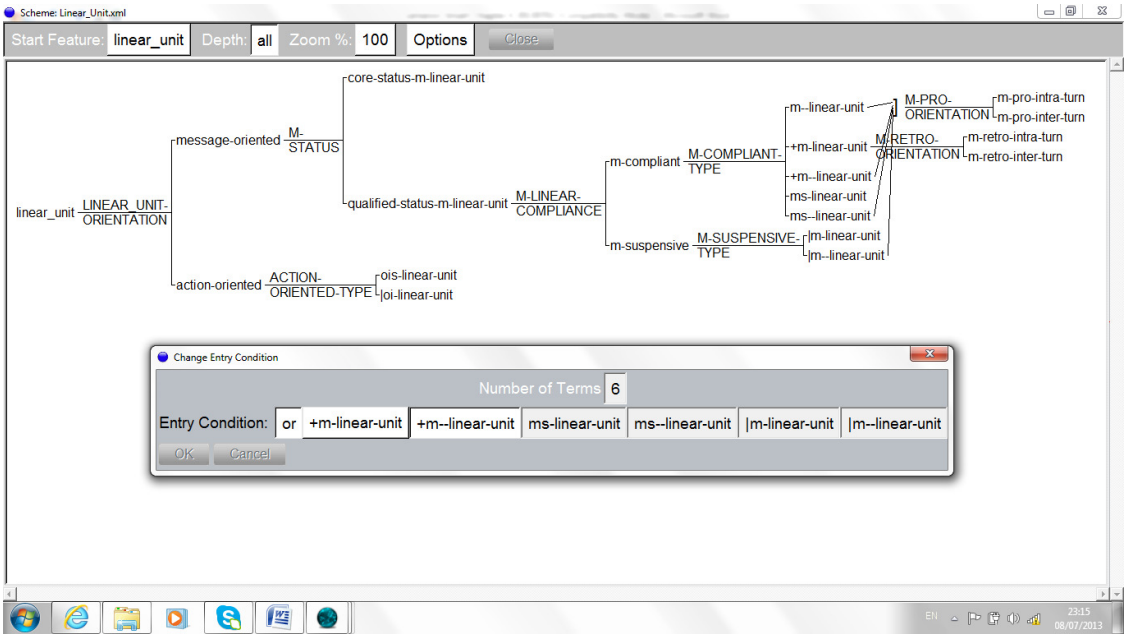


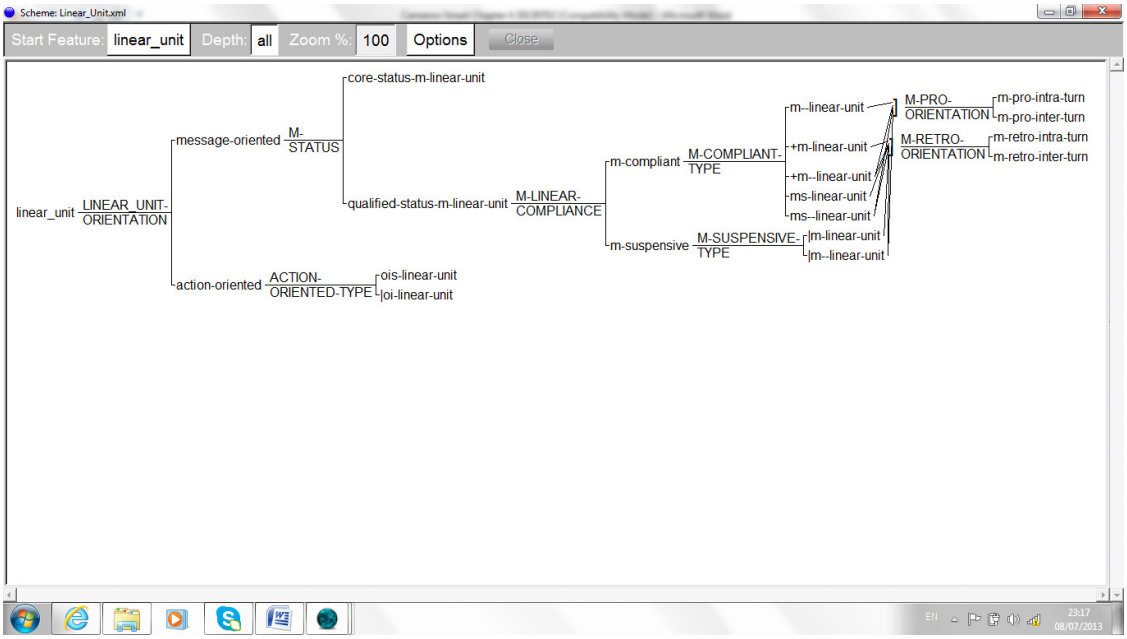
Step 6: Creating retro turn orientation for M linear units

i. Add system to *+m linear unit*. Rename the system: *M RETRO ORIENTATION*.

ii. Rename features: *m retro intra turn* and *m retro inter turn.*

iii. Change Entry Condition of M RETRO ORIENTATION system: Number of terms: 6; Entry condition: *or +m linear unit*, *+m– linear unit*, *ms linear unit*, *ms– linear unit*, *|m linear unit* and *|m– linear unit*.





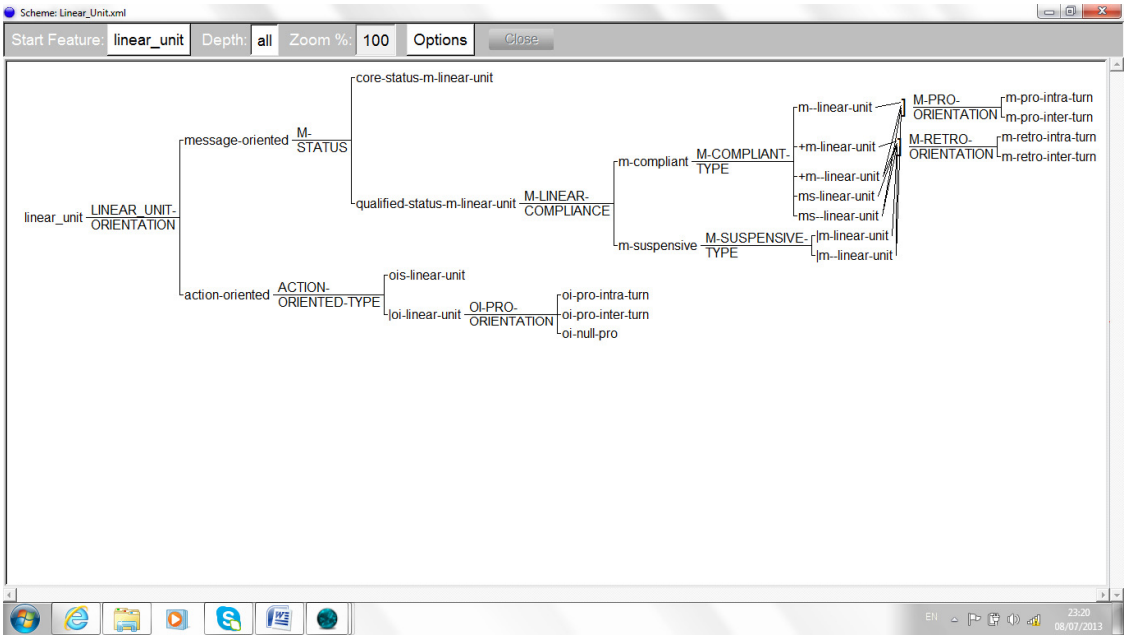
Step 7: Creating subcategories of OI linear units

i. Add system to *action oriented*.

ii. Rename the features: *ois linear unit* and |*oi linear unit.*

iii. Add system to |oi linear unit. Rename the system: OI PRO ORIENTATION.

iv. Rename features: *oi pro intra turn*, *oi pro inter turn* and *oi pro null.*[125]



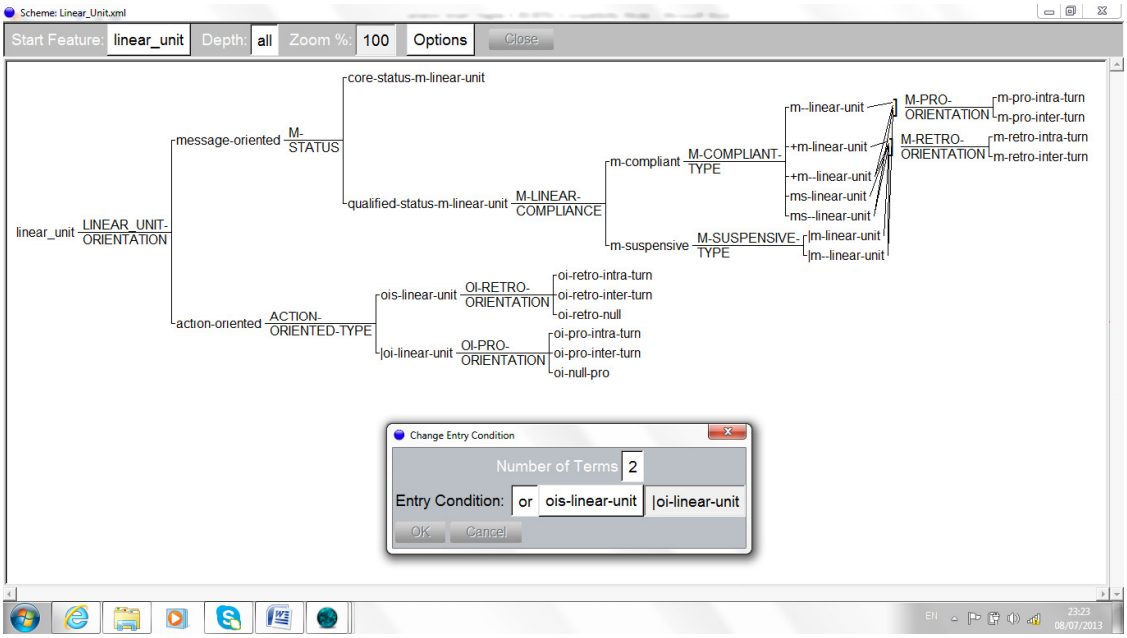
Step 8: Creating retro turn orientation for OI linear units

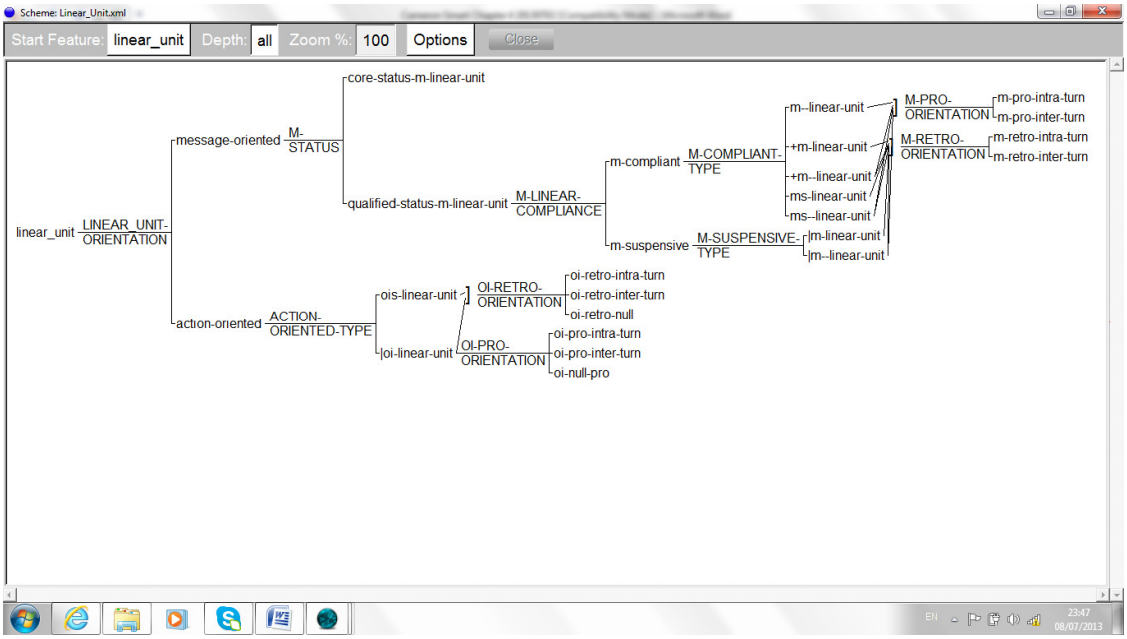
i. Add system to ois linear unit. Rename the system: OI RETRO ORIENTATION.

ii. Rename the feature: *oi retro intra turn, oi retro inter turn* and *oi retro null.*

iii. Change Entry Condition of OI RETRO ORIENTATION system: Number of terms: 2; Entry condition: *or ois linear unit, \oi linear unit*.

[125] There is a need to include an *O retro null* and *O pro null* option for OI linear units. *O retro null* are |OI linear units which are suspensive with no link to a previous linear unit, such as salutation |OI linear units, e.g. *Hi!* In contrast, *O pro null linear units* are those which are |OI linear units but which do not prospect anything. These are generally short phrase acknowledgement responses e.g. *Ha!* or *Shut up.*

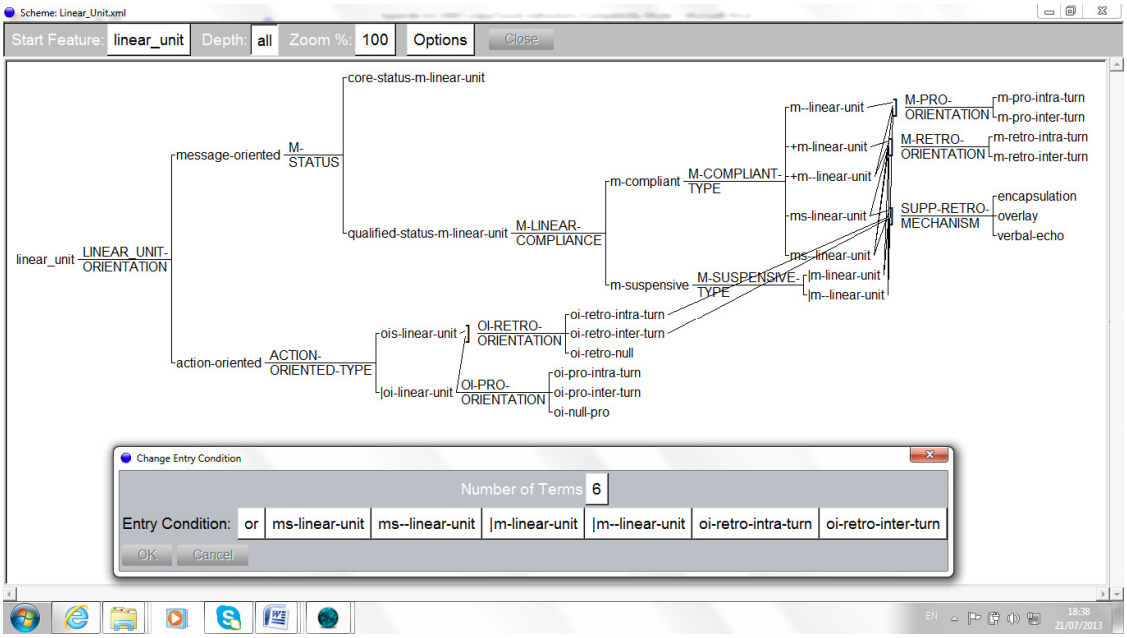




Step 9: Creating categories of supplementary retrospective mechanisms

i. Add system to MS linear unit. Rename the system: SUPP RETRO MECHANISM.

ii. Rename and add features: *encapsulation*, *overlay* and *verbal echo*



iii. Change Entry Condition of SUPP RETRO MECHANISM system: Number of terms: 6; Entry condition: *or ms linear unit*, *ms– linear unit*, *|m linear unit*, *|m– linear unit*, *oi retro intra turn* and *oi retro inter turn*