Structuring Information in Written English: A Re-appraisal of the Systemic Functional Approach to Information Structure

Thesis submitted in accordance with the requirements of
the University of Liverpool for the degree of
Doctor in Philosophy
by Nicolas Antony James Moore

October 2010

Abstract

The concept of Information Structure in English originated in Systemic Functional Linguistics, and has developed to function as *newsworthy* information, or what the speaker directs the listener to focus on. In spoken English, Information Structure is realised in intonation, with the tonic foot identifying New information. In written English, consensus has gradually converged on the same function being realised by final position in a clause. While this is assumed to be the case, there has been no study so far that demonstrates a role for Information Structure in written English independent of Reference or Theme. The aim of the present study is to demonstrate the existence and role of an independent Information Structure in written English, within the context of a modern Systemic Functional theory. The main strategy employed is to investigate the systems in the Textual Metafunction that are realised within clauses by analysing texts for the systems of Participant Identification and Tracking, and for Theme and Rheme, using a largely quantitative approach on a range of texts from a single register.

Systemic Functional theories of reference are compared with other perspectives, resulting in a re-evaluation of the theory of *bridging* to include an integration of a taxonomy of sense relations. Quantitative analysis of a small corpus of texts reveals significant patterns for Participant Identification and Tracking. Starting with a clarification of Theme within a Systemic Functional perspective, text analysis describes the interplay of Theme with Participant Identification and Tracking to quantitatively demonstrate the unmarked correlation of Theme with Presuming and Rheme with Presenting reference.

An investigation of the development of the concept of Information Structure both within and beyond SFL concludes that sequence is likely to realise both Theme and Information in written English. This hypothesis is tested against the corpus of texts used in this study. By examining the interaction of the three systems of Participants, Theme and Information, quantitative and discoursal patterns emerge that reveal an independent function of Information Structure. Finally, functional, psychological, neurological and historical evidence is examined to explain the importance of sequence in realising Information Structure. It is argued that spaces and punctuation marks realise Information Structure in written English. The implications for theories of linguistics are explored.

Table of Contents

List of Tables

List of Figures

Acknowledgements

With all my heart I gratefully acknowledge the valuable contributions of:

- Geoff Thompson to whom I am deeply grateful. His guidance has been invaluable throughout this long, arduous process. He has ALWAYS (and I mean, always) been available for discussion, comment and feedback. It goes without saying that the current work would have not materialised without his constant support, but I also appreciate just how lucky I am to have such a caring and committed professional as a supervisor, as well as a warm and thoughtful person as my friend in Liverpool.
- Michael Hoey; who has always opened his door to me for inspirational, provocative and valuable discussion, even when his many other commitments place great demands on his time. Many of the ideas in this thesis derive from these discussions, although he is not to be held responsible for the direction I may have taken them.
- The SFL Community; at conferences and online for their openness to discuss issues, in particular to Michael Halliday for always having the time to stop and chat about any point of interest and for encouraging SFL scholars to build theory through cooperation rather than conflict, and to Jim Martin, Ruqaiya Hasan and Christian Matthiessen and others for continuing to develop SFL in wider directions.
- Carys Jones, fellow organisers, and participants at ESFW05 and Eija Ventola; for encouragement, comments and critical feedback in the preparation of the presentation and paper that details some of the ideas contained in chapter 2.
- Martin Davies; who offered inspiration and support in correspondence from the very start of (and even before) this study.
- Peter Fries; for the generosity with which he is willing to share time and ideas, without notice and without any prejudice.
- Mick O'Donnell; for both Systemic Coder and for the UAM Corpus Tools, for timely and valuable feedback on the software, and for encouragement, suggestions and considerable effort to improve the software by listening to the users.
- Mike Scott; for encouragement, valuable advice and Wordsmith Tools.
- Ahmed Bentiba; for valuable expertise in solving a major software problem.
- Chris Butler; for comments and advice far beyond requirements to produce a superior thesis.
- Phyllis Burns and Eileen O'Brien; colleagues, office mates and friends for constant encouragement, support and friendship.
- David Vernon; for motivation through reminding, nagging and cajoling at the most opportune moments.
- My parents Ray and Diane and my brothers Stuart and Nigel; for the support and belief they have all shown in me.
- Finally, my wife, Özlem, and my children, Marcus and Elif, who have been without their husband and father for too many summers and weekends, who have been quiet around the house when necessary, who have put up with me disappearing to my 'office' when we should be doing other things together, and who have helped me in so many ways to complete this project.
  I dedicate this thesis to them for the love, support and inspiration without which I would never have started on this path. It is for them that I endeavour.

Chapter 1 Introduction: In Search of Information

1.0 Introduction

This introductory section attempts to establish, in broad terms, the direction, focus and major concerns of the present study. It outlines the general aims and guiding principles of the whole study, describes the Textual Metafunction in Systemic Functional Linguistics (SFL) – a major concept in the study – and the methodology of the study, and offers a brief outline of the organisation of the remaining chapters.

1.1 Aims

The main aim of this study is to demonstrate how information is structured within the clause in the Systemic Functional model of written English. A thorough investigation of this area in written English should contribute to a better understanding of information structure both within Systemic Functional Linguistics (SFL) and beyond. SFL distinguishes systems of meaning-making that are often treated together in other theories of information structure. The next section (1.1.1) will briefly introduce some of the different theories of structuring information to be compared to the SFL model in the study. In SFL, information structure is part of the Textual metafunction, and so section 1.1.2 describes the role of the Textual Metafunction.

It is through the main systems in the Textual metafunction of Participant Identification and Tracking, Theme, and Information that SFL treats those features of language which are often referred to as Information Structure in other approaches. This study will investigate these textual systems in detail, through discussion and both quantitative and descriptive textual analysis. It will also detail how the three systems interact to create meaning.

The concept of information structure originates in Systemic Functional Linguistics. This study briefly discusses the development of the concept, but also demonstrates how other approaches regard information structure. The review in this study will attempt to show where the SFL approach has developed comprehensive descriptive tools for these concepts so that they can be applied to discourse analysis, with the interaction of the separate systems creating a meaning-making dynamic. However, the study will also

attempt to identify where possible innovations from other theories may benefit and develop SFL theory. The study ultimately attempts to demonstrate how the systems of Participant Identification and Participant Tracking (chapter 2), Theme (chapter 3), and Information (chapter 4) contribute meaning to the written clause individually and in combination.

As a consequence of focussing on written English, a further aim evolves. SFL typically uses different realisations to assign New information in written English and in spoken English, but there has been no attempt to explain why this may be the case. This study aims to establish why a realisation of New information in written English has evolved separately from that in spoken English. Chapter 5 describes historical and neurological factors that may have influenced this development.

1.1.1 Structuring Information

In speech or writing, one function of language is to offer information (in the 'folk' sense) to a listener or reader. At the same time, there is a lot of redundancy in discourse: repetition, contextualisation and 'phatic communion' (discourse with the primary purpose of furthering social relationships) contribute significant amounts of discourse but little new knowledge to our spoken/written production. It is to the advantage of the speaker/writer and listener/reader if information that is considered central can be distinguished from information that may be ignored without disrupting the flow of discourse. Typically, the information that we provide can be divided into information that we want our audience to know and information that is peripheral; information can be 'structured'.

Information structure has come to mean many things in linguistics. Most contemporary theories of structuring information can trace the concepts of information structure and New information directly or indirectly to Halliday's (1967a) seminal paper (a claim which is revisited throughout this thesis). Halliday (1967a) disconnected Theme and Information from Mathesius' Prague School notion of Theme. At about the same time, Firbas (*e.g.* 1968) was also contributing the idea of new information to the Prague School theory that he would develop into Communicative Dynamism (1992), and Bolinger (1958; 1965) was investigating the link between intonation, information and word order. Halliday proposed that speakers divide discourse into units, realised naturally by a tone unit, and use stress and intonation, specifically the tonic foot, to distinguish what they consider important.

Halliday ascribed the term Information structure to the function of dividing discourse by tone into manageable units. Within Information structure, the unit of information is divided into New and Given. The tonic foot realises New information, and any remaining unstressed parts of the tone unit are labelled Given. Information, New and Given in this sense are technical terms and are distinct from the folk sense of information meaning previously unknown; in a Systemic Functional approach, information structure is narrowly defined to function to divide text into manageable units of discourse, and New Information functions to direct a listener's attention to what the speaker considers the most important (Halliday, 1967a; 1967b; 1976; Halliday and Matthiessen, 2004). Chapter 4 in this thesis investigates the SFL notion of information structure in written English.

Related to the folk notion of new information are the SFL concepts of Theme and Participant Identification and Tracking. Within a clause, there is a sense that the clause is 'about' something – that it provides information about a topic. In some grammatical theories, this structure is labelled Topic-Comment. The related structure in SFL theory is Theme-Rheme. Chapter 3 discusses the main issues related to Theme-Rheme structure. Finally, new information in a clause, in the non-technical sense, can be revealed through the use of determiners, with indefinite articles being the typical indicator of something "new" in the context. While some theories equate the system of determiners with "new" information, as described in chapter 2, the SFL concepts of Participant Identification and Participant Tracking recognise that determiners only apply to nominal groups (and similar) and so do not apply to as broad a range of linguistic features as required by the technically-defined intonational-based description of information structure. Chapter 2 analyses the role of Participant Identification and Tracking, particularly the functions of Presenting and Presuming reference, as distinct from the notion of new information.

The concepts of Given and New Information and Information Structure have been developed in various directions both in SFL theory and in other theories. Notably, Chafe (1970), Clark (1977) and Prince (1981) tie the concepts of Given and New to psychological and pragmatic accounts of what is happening in the minds of interlocutors, suggesting that certain referents may be considered New or Given in the discourse context and accorded informational status through tone and/or reference. These contributions are reviewed in chapter 2, along with computational approaches to reference, especially in the field of anaphora resolution. Another structure that is often related to either the realisation of information structure, particularly sequence in English, or to the function of emphasising a part of the clause is the Topic-Comment structure.

Topic-Comment is subsumed within SFL theory in the system of Theme-Rheme, but very few studies outside SFL and Prague School linguistics have adopted the concept of Theme. In chapter 3 we review developments in Theme and describe similarities between Topic-Comment and Theme-Rheme. Steedman (1991), Lambrecht (1994) and Jackendoff (2002) incorporate the concept of information into broader syntactic theories, combining psychological states, word order and intonation. These are further discussed and compared with SFL approach in chapter 4. Functional studies of information structure have also developed the theory beyond Halliday's original contribution (see Butler 2003; 2005 for review). Throughout this study, developments in related theories are reviewed to reveal where they may be able to contribute to the SFL approach and to identify where the SFL approach may be considered more effective.

1.1.2 The Textual Metafunction

Within Systemic Functional Linguistics (SFL), language is theorised to have evolved into three configurations of meaning associated with three modes of realisation. The resources for making meaning in a language are assumed to function in these three strands simultaneously. These functions are called Metafunctions because they create a level of organisation generalisable across a wide range of language features. Meaning-making resources function concurrently to construe reality, to enact relationships and to contextualise discourse. The three Metafunctions are, respectively, the Experiential (divided into Ideational and Logical), the Interpersonal and the Textual.

The Experiential Metafunction construes a reality based on categories of experience and helps each of us to make sense of the world around us (Halliday and Matthiessen, 1999). The Experiential Metafunction is divided into the Ideational Metafunction, which is typically realised by constituent structures (Halliday, 1979), and the Logical Metafunction, typically realised by iterative structures (Halliday and Matthiessen, 2004 p.61). The Ideational metafunction allows reality to be construed through language, and is realised in grammatical systems such as transitivity. The Logical Metafunction construes relationships between stretches of discourse, clauses, groups and lexical items (see sections 3.6, 3.7.2 and 3.8.9.1), and may be recognised most readily in conjunctions. The Experiential Metafunction has been the focus of the majority of linguistic theories. This partly explains why many systemic functional analyses, concepts and theories are reinterpreted by other linguistic theories in terms of the Experiential Metafunction (see chapters 3 and 4 for examples.)

The Interpersonal Metafunction enacts social relationships (Halliday and Matthiessen, 2004 p.61). Language functions to establish, create, recreate and enforce social roles and relationships within a community. Interpersonal aspects of language include 'speech acts', such as demanding or offering information or goods and services, and are typically realised by prosody and by 'pulses' through the clause (Halliday, 1979). In a number of linguistic theories, these aspects of language are often assigned to 'Pragmatics' or 'Speech Act Theory' and thus not considered a part of syntax or grammar as such (see Thibault and van Leeuwen, 1996 for a critique of this approach). Recently, there has been a great deal of interest in SFL in the Interpersonal resources that comprise 'Evaluation' (Hunston and Thompson, 2000) and 'Appraisal' (Martin and White, 2005).

The function of the Textual Metafunction is to create "relevance to context" (Halliday and Matthiessen, 2004, p.61), and is characterised by wave-like patterns of prominence (Halliday, 1979). It draws on a range of systems to achieve this: Conjunction, Information, Theme, Determination and Cohesion. Although most systems of the Textual Metafunction can be identified within a clause, they function within the contexts of discourse (co-text), situation (context) and culture (socio-historical moment). Thus, this study will draw heavily on the Systemic Functional Linguistic theory that takes discourse as its basic unit of grammatical analysis, namely that set out in Martin's (1992) *English Text*. The terms adopted by Martin (1992) represent the functions associated with discourse and are intended to complement Halliday's (1985; 1994; Halliday and Matthiessen 2004) clause-based grammar. However, systems in the Textual Metafunction that require more than one clause to function, such as Conjunction (the combination of clauses – traditionally, simple sentences – into clause complexes – or complex sentences) and Activity Sequence (the Logical ordering of clauses) (see Martin, 1992), will not be investigated here.

Matthiessen (1992) characterises the Textual metafunction as a second order system. It is second order to Experiential and Interpersonal meanings neither in terms of sequence of realisation nor in importance, but is second order in that it uses Interpersonal and Experiential meanings to create meaning that is internal to the discourse itself. The Textual metafunction takes Experiential and Interpersonal meanings and as it ties them to a particular context it also adds prominence of various kinds. Thus, certain meanings are given relative value through the systems in the Textual metafunction, as neither the Experiential nor the Interpersonal metafunction assigns different values to their

realisations (Matthiessen 1992). It is only through the processes of both placing Experiential and Interpersonal meanings in context and giving them value that Experiential and Interpersonal meanings can operate in context. Consequently, the Textual metafunction *instantiates* Experiential and Interpersonal meanings.

As the following chapters testify, there is considerable variation between theories of the Textual systems of Reference, Theme, and Information. Different linguistic theories use the terms and concepts in differing, often contrary, ways with many terms overlapping and substituting for each other. Even within a school there can be disagreement over some definitions. Bearing this and the preceding discussion in mind, I will use Fries (*e.g.* 2000) to act as a guide through these different terms and systems. This is partly because Fries has discussed the textual metafunction at length, but also because he has consistently managed to prise apart the different functions and their realisations; for instance Fries (1981) distinguished the 'combining' and 'separating' approaches to Theme.

Fries (1992; 2000; 2002) discusses the different systems that operate within the clause to realise Textual functions. He distinguishes the systems of Presenting and Presuming Reference, Theme and Rheme, and Given and New Information. Specifically, Fries (2000) argues that the system of Presenting and Presuming Reference and the system of Given and New Information are "*two independent dimensions*. I have not argued, but will happily admit, that there is a strong correlation between the two dimensions." (p.103, original emphasis). That is, the realisation and the meanings associated with the two systems are independent and contribute their own function to a text. As Fries points out, the two systems are typically correlated, producing an unmarked relationship between Presuming Reference with Given and Presenting Reference with New Information, but we should never confuse an unmarked relationship with a definition. Significantly, and in contrast to other approaches, choices of Reference and Information are independent of an external reality: "The externally observable facts do not *determine* either information focus or the use of presenting vs. presuming reference." (Fries, 2000, p.103, original emphasis) That is, the resources used in these systems are chosen by the speaker or writer to create Textual meaning – to contextualise participants and other aspects of the clause according to the meaning required. This study into systems in the Textual metafunction will be guided by the distinctions noted by Fries and will draw on the analytical framework in Martin (1992) – the origin of the distinction between Presenting and Presuming Reference.

1.2 Guiding Principles

Throughout this investigation, certain guiding principles will be used to evaluate different perspectives and analyses. These are based mainly on Systemic Functional Linguistics and a corresponding embodied and constructivist or phenomenological approach.

1.2.1 Systemic Functional Linguistics

This is a study of the three clausal systems in the textual metafunction from the perspective of Systemic Functional Linguistics (SFL). One of these systems is known in SFL as Information Structure. This study reveals that the linguistic features identified as information structure in other schools of linguistics correspond with at least one of these three textual systems. Other schools of linguistics, as well as references to psychology and philosophy, will be evaluated in order to establish the role and interaction of Participant Identification and Tracking, Theme and Information in making meaning in the textual metafunction in written English.

Within Systemic Functional Linguistics there are different schools, and a range of criteria for arriving at some analyses co-exist. In order to allow differing perspectives to collaborate rather than compete, an analogy with language itself has been developed – that of dialects of a language. Perhaps the most dominant dialect within SFL is the "Sydney dialect", which itself contains variants. Matthiessen and Martin are two theorists working within this dialect. Of the other dialects, one obvious candidate is the Cardiff school. Based in Cardiff, Fawcett (*e.g.* 2000) adheres to the same principles of a Systemic Functional approach but, partly as a result of focussing on a computationally-generative version of the grammar, often produces different analytical models of English. Fawcett's colleagues Tench and Tucker have also published within this dialect. This study will attempt to incorporate different dialects of SFL, as well as the dialects, "accents" and "languages" of other linguistic theories. Ultimately, though, the study will mainly evaluate contributions to the study from the perspective of the Sydney school, and particularly Martin's (1992) Discourse Semantics – an "accent" within the Sydney dialect.

1.2.2 Dynamic and Synoptic Perspectives

A distinction is often made in SFL between dynamic and synoptic analyses. This is the difference between, respectively, an analysis of text as it is revealed word-by-word to the hearer or reader as a process, 'on the fly', or *in vivo*, as opposed to an analysis of the whole text after it has been produced, as a product, as a whole text, *in vitro*. O'Donnell (1990, 1999) and Ravelli (1995), in particular, have raised the issue of the two complementary perspectives, and Martin (1992) often refers to dynamic and synoptic analyses. A recent study that emphasises the dynamic approach – one that is sympathetic to an SFL approach – is Sinclair and Mauranen's 'Linear Unit Grammar' (2006) which investigates the syntagmatic units of English from the perspective of a listener or reader. This study will recognise that dynamic and synoptic analyses are important and that any account of information structure will need to take both into account.

1.2.3 Degree of Delicacy

One reason for attempting to improve current SFL descriptions of information structure is that there is a major discrepancy between the precision with which New information can be identified in spoken and in written English. The boundary between New and Given information (if Given is present) in speech is inherently indistinct, but the focus of New information is identifiable down to the syllable in spoken English because it is realised by the tonic foot. In written English, however, even the best descriptions (*e.g.* Fries' (1992) N-Rheme and Matthiessen's (1995a) Culmination) tend to offer descriptions of information focus that can extend over a large proportion of the clause; the definitions are not very focussed. Although the transition from Given to New may not be easy to identify, and taking into consideration the wave-like structure of features in the Textual metafunction, it should be possible to identify more clearly the location of the syllable, word or group that realises the focus of New information in written English. Consequently, most of this study works at the level of Group, on the assumption that the experiential Thing of the relevant group is most likely to be the focus of New information (see section 2.5).

1.2.4 SFL and Constructivism

One area in which SFL theory has developed significantly in recent years is in the connections that are being made to brain sciences and post-cognitive psychology. The approaches to the structuring of information in this study, including the three systems in

the textual metafunction, will be evaluated by the extent to which they correspond with the latest developments in our understanding of the way that we learn and use language from both an *inter-organism* perspective – the social context – and from the *intra-organism* perspective – how language develops and operates within the brain. In this section I will attempt to outline the developments in our understanding of the *languaging* brain, in biological and cognitive sciences, as well as corresponding developments in SFL theory.

1.2.4.1 Post-Cognitivist Psychology

Although SFL has rarely taken an explicitly psychological approach to language processing (Butler, 2008), it has often interacted with related disciplines (*e.g.* Halliday, 1995; Matthiessen, 1995b; Benson and Greaves, 2005; Thibault, 2005). More recently, its constructivist approach has been matched with developments in neuroanatomy and the biological sciences. The significant breakthrough of 'biology of cognition' and autopoiesis (Maturana and Varela, 1987) has developed to the point where, I believe, the arguments are so convincing that psychology as a discipline can no longer ignore the physical and biological roots of consciousness.

Maturana and Varela's theory of autopoiesis describes how the evolutionary imperative of multi-cellular creatures is only observable from the outside, and any attempt at explanation inevitably involves the observer's perspective, epitomised by the maxim "Everything said is said by someone" (1987, p.27) – that is, there is no observation without an observer whose subjective position is a necessary part of the observation process. From this phenomenological standpoint they proceed to explain how consciousness, and consequently language and ethics, is the result of the organism's interaction with its environment – its 'structural coupling'. Maturana and Varela (1987) stress that any theory of the 'mind' must be based in physical reality – that a metaphysical separation of the mind from the brain is unnecessary – and specifically discuss language within their biological framework. They conclude that communication is ill-served by the Message-Sender-Medium-Receiver-Message model, noting that the only "truth" of a message is its effect on the receiver, regardless of an observer's presumption of content or intent.

Within an autopoietic theory, the demand for a "nonrepresentationist view of knowledge based on the sense-making capacity of an autonomous living system" (Maturana and Varela, 1987 p.254) entails not a realist denial of external reality but a requirement on the

part of a living being to be actively involved in the interactions with the environment that will determine which real-life distinctions are valued and become part of the organism's ontogenetic history of coupling, and, therefore, their system of values and meaning. Clearly, one part of a human's environment, influencing events of structural coupling, is their integration into social units. In terms of developing language, Maturana (1978) refers to this as the 'consensual domain', and it is this part of the environment that allows for the continuity and reproduction of a rich diversity of linguistic behaviour. Maturana is explicit in rejecting the nativist theory of language acquisition:

> the superficial syntactic structure of a given natural language can only be a description of the regularities in the concatenation of the elements of the consensual behavior. In principle, the superficial syntax can be any, because its determination is contingent on the history of consensual coupling, and is not a necessary result of any physiology. Conversely, the 'universal grammar' of which linguists speak as the necessary set of underlying rules common to all human natural languages can refer only to the universality of the process of recursive structural coupling that takes place in humans through the recursive application of the components of a consensual domain without the consensual domain. (Maturana, 1978 p.52)

That is, from a biological perspective, while contexts of language use vary across history and culture, what is common to human language is the process of coupling with the natural and social environment. Maturana explicitly rejects a generativist programme of linguistic description: "Superficial and deep syntactic structures are features of one description of linguistic utterances, not of the process of their generation." (1978, p.53) Steels (1998, 2000) has demonstrated that, in an environment that demands cooperation for survival, communication and language will emerge from agents endowed with only general learning skills, and so there is no need for any genetic coding prior to language learning. The implications of autopoiesis for theories of language are further explored by Kravchenko (*e.g.* 2002, 2006).

The separation of a metaphysical mind from the physical brain is also rejected by leading neuroscientists. Theories of consciousness have been developed that attempt to identify the physical properties and processes of the brain that enable consciousness. It would appear that consciousness is made possible by the recursive properties of the brain. For Arbib (*e.g.* 2000), this is partly achieved through 'mirror neurons' which, because of their ability to imitate and simulate, enable learning (also see Rizzolatti and Craighero 2004). Mirror neurons empathetically simulate the motor processes involved in an observed (or heard) action; as you carry out a physical action my brain attempts to implement the same neural programmes as yours at the speed of neurochemical transfer (which far outstrips

any physical action), although it does not implement them physically. This allows the observer to learn and improve, even before attempting to copy modelled behaviour. Recent research strongly suggests that the brain carries out the same processes with linguistic behaviours – as you speak, I am carrying out the neural processes required to say the same thing (D'Ausillio *et al.* 2009; Devlin and Aydelott, 2009). Edelman (1999, 2004) emphasises the fantastically extravagant structure of the brain, and focuses on the multitude of 're-entrant' systems that distinguish the human brain from other animals. Edelman suggests that through re-entry the human brain achieves consciousness as experiences can be simulated by re-activating the same neural pathways as were stimulated during the first experience. Consequently, learning, rehearsal, prediction and various other processes of consciousness can all be explained by features of the multitude of re-entrant systems.

In short, there is no evidence for a reality that exists without interpretation by the biological systems of the organism that are constantly in interaction with both the environment, including the social world, and the central nervous system. It is the patterning of interactions with the outside world that allows for interpretation. The implication of these studies is that since the brain does not need a separate mind, psychological theories that presume a mind must be re-examined, particularly those that rest on a representationist paradigm which demands meanings that exist prior to interpretation, as these have no basis in observation of biological and physiological systems.

Rejoining the mind with its physical brain has resulted in another significant development in psychology – embodied cognition (*e.g.* Barsalou 2008) – which has had a significant influence on Artificial Intelligence (Froese and Ziemke, 2009; Vernon and Furlong, 2007). This movement represents a re-emphasis on the physical nature of learning; cognition and consciousness will always take place within a physical body within a physical and social context. As Maturana and Varela put it "All knowing is doing and all doing is knowing" (1987 p.27) – knowing and learning are ultimately physiological processes brought about through experience. Approaches to language learning in infants that emphasise the role of the embodied experience of learning have become increasingly popular (and may also be applicable to second language learners), partly because of the debate surrounding the *poverty-of-stimulus* (see Sampson, 1997; Pullum and Scholz, 2002 for typical critiques). A re-examination of the central role played by the interaction of the developing language user

with the physical and social environment has supported the development of these new theories.

Emergence, connectionism and embodiment are developments in the biological and psychological sciences that have been closely reflected in developments in SFL, which is inevitable considering the phenomenological and social-constructivist approach evident in the work by Halliday (*e.g.* 1975) and Hasan (*e.g.* 1986/1996) on language learning in infants. By attributing a proto-language to infants, in contrast to the typical view that infants' only goal is to copy adult language, and by charting the development of language through the different stages of protolanguage, generalization, abstractness and metaphor (Halliday, 1993), Halliday opposes the typical approach to language learning epitomised by the term *acquisition*:

> it is perhaps not too far-fetched to recognise in the use of the term *acquisition*, a further implication that structure, and therefore language itself, is a commodity of some kind that the child has to gain possession of in the course of maturation. (Halliday 1975, p.1)

While many theories, in an attempt to incorporate an embodied perspective, have had to return context to the phenomenon and study of language, SFL has always attempted to integrate the context of culture, genre, register and co-text into its descriptions of meaning. Thibault (2004a, 2004b) and Lemke (2000) have been particularly active in integrating biocognitive accounts and ecosocial semiotic theory, respectively, with SFL descriptions.

1.2.4.2 Construing Experience

SFL is a multifunctional theory (see section 1.1.2). It aims to link a grammatical account with all aspects of language use – including aspects that in other theories may be separated from grammar into semantic or pragmatic descriptions. The Textual and Interpersonal metafunctions account for grammatical choices that 'engender discourse' and 'construct our social relationships', respectively (Halliday and Matthiessen, 1999). One description of an SFL 'construal' perspective on the Ideational metafunction pinpoints the narrow view of language engendered by a truth semantics-based approach (Thibault, 1999). Freeing the theory from a preoccupation with truth-based semantics and symbolic logic allows SFL to develop a theory of Ideational grammar that does not require a representationist perspective.

SFL joins post-cognitivist psychology, and some linguistic theories, in problematising a representationist approach to language. In assuming that natural languages represent the world by providing structures and lexis to describe a pre-ordained order, the role of humans is removed from the cognition required to make sense of our environment. This is precisely the point that embodied approaches to learning take as a common critique against prior theories of symbolic representation, in both linguistics and AI. Meaning in a connectionist or embodied theory does not reside in symbols of an external reality, since that reality is defined only in terms of the individual's history of interactions with reality, with no necessary split between the internal and external. Rather, meaning "resides in complex patterns of activity that emerge from the interactions of many such constituents [of experience]" (Varela, Thompson and Rosch 1991, p.100) and are significant only in relation to their contribution to action, including *languaging*.

If connectionist and emergent models of learning have validity they can be used in support of models of language learning based on usage. Proposals for the 'adequacy of the stimulus' are being put forward, arguing for models that emphasise usage (Bybee 2006), probability (Chater and Manning 2005), connectionism (Westermann, Ruh and Plunkett 2009) and emergence (MacWhinney 1998, Behrens 2009). While there are clear differences between all of these proposals, they are united, with SFL, in assuming that normal human learning processes are sufficient for language learning. It is not always clear, however, to what extent each of these approaches rejects the demand for representationism.

Perhaps the most significant statement of a non-representational description of Experiential grammar is Halliday and Matthiessen (1999). Instead of regarding Experiential grammar as a representation of either an extra-linguistic material reality or a non-linguistic mental reality, Halliday and Matthiessen (1999) stress that meaning is at the base of human experience; meaning allows us to order experience. This meaning is not inherent in the world, as a representational theory would have us believe. Meaning is constructed in social reality, and Ideational meaning is characterised by socially-sanctioned systems of categorisation:

> Categorizing is a creative act: it transforms our experience into meaning, and this means imposing a categorical order rather than putting labels on an order that is already there … what our semantic resources enable us to do is to construe those analogies which yield categories resonating with what as a species, and as members of a particular culture, we have found to carry material and symbolic value. (p.68)

This bold statement aligns SFL theory with a range of new theories about the way that the brain works based not on conjecture about psychological processes, but derived from studies of the structure and real-time functioning of the brain. Halliday and Matthiessen (1999) make explicit their constructivist approach by drawing on Hjelmslev, who has long inspired the theoretical foundation of SFL:

> The view we are taking is a constructivist one, familiar from European linguistics in the work of Hjelmslev and Firth. According to this view, it is the grammar itself that constructs for us our world of events and objects. … Meanings do not 'exist' before the wordings that realize them. They are formed out of the impact between our consciousness and its environment. (p.17)

Although Halliday and Matthiessen (1999) are specifically modelling Experiential grammar, one main guiding principle of the current study will be to ensure that the conclusions are consistent with a constructivist approach to the grammar, providing the same approach in the Textual metafunction as offered in the experiential metafunction. Therefore, the theory developed will need to consider how the contextualising features of the Textual metafunction enable meaning to be construed within a context.

This apparent digression into a cursory glance at developments in neurobiology is required in order to set the direction that any psychologically-related conclusions must follow. That is, not only must this study be compatible with a constructivist approach in linguistics, but psychological processes, including those presumed for reading (see especially chapter 5) or for tracking participants through text (chapter 2), must also be based on physical descriptions of the brain structure rather than unfalsifiable theories of psychology that support such contentious concepts as 'the poverty of stimulus' (*cf.* Pullum and Scholz, 2002), autonomous syntax (*cf.* Thibault, 1999) or a language acquisition device – none of which have been demonstrated to exist in the brain. This study is thus part of an attempt to promote a linguistics and a psychology that is compatible with a model of cognition that is constructivist, embodied and based on neurological research.

1.3 The Study

Developments in the main areas of study in the textual metafunction – participant Identification and Tracking, Theme, and Information – will be discussed, and comparisons with other theories with similar concepts will be made. This study also employs a text-based approach to analysis which allows for an examination of the

interaction of the three systems in text. The methods and the corpus of texts are outlined in this section.

1.3.1 Methods

The methods used most extensively in this study are a combination of a review of previous studies and text analysis for the three main systems of the Textual metafunction: Participant Identification and Tracking, Theme-Rheme, and Given-New Information. Initial text analysis is quantitative in nature, and is followed by discussion. A discussion distinguishing the approach taken to reference in SFL is followed by Participant Identification and Participant Tracking analyses. Similarly, a discussion of the main areas of controversy in the system of Theme precedes Theme-Rheme analysis. The two systems of participant Identification and Tracking are then analysed in relation to Theme. A discussion of Information structure, paying particular attention to how it differs from the other systems in SFL theory and where this coincides with other theories, is followed by a text analysis of Given-New information, and then its interactions with the other systems. This leads us to suggesting a model of Information Structure in written English, which is supported by evidence from a range of sources in the final chapter.

1.3.2 The Corpus

The texts that are analysed in this study were produced within a particular social context. They are all examples of a particular Register, as defined by their Tenor, Mode and Field (see Table 1.1). Each one is pedagogic text – it is Expository: Analytical (Martin 1989) – and so the Tenor of the text is an unequal one. The Tenor can be characterised as an explanation by someone – an expert – who has all of the knowledge for someone who has none – a novice. The Mode for each text is written English in a published textbook chosen as the main reference for an undergraduate engineering course. Each text comes from a corpus of the texts at Khalifa University (formerly Etisalat University College) in the UAE, where I am a practising teacher of English for Academic Purposes. The students at this university are all native Arabic speakers who have to deal with these English-language texts in their degree programme, which is accredited by the Institution of Engineering and Technology (IET) in the UK. The Field for these texts, therefore, is engineering, specifically communications, computer or electronics engineering at undergraduate level. The only exception to this general description is a text that varies in its Field, as it comes from aeronautical engineering. This was included to identify the

variation in the Textual metafunction dependent on Field. It is hoped that subsequent to this study specific advice in reading strategies to students at Khalifa University can be offered, as well as more general advice to other engineering students.

FIELD	MODE	TENOR
Technical; Related to engineering and related sciences; Taxonomic, expository	Printed media; Published commercial material; Written text and accompanying graphical aids	Asymmetric access to knowledge; Tutorial by expert for novice with limited background knowledge in science; Instructional

Table 1.1 Register features of corpus of texts for analysis

In total, the nine texts total more than 10,000 words of analysis across four systems of analysis. Table 1.2 (over) provides bibliographical and quantitative details on all of the texts. The shortest text 'AN' was only 385 words, while the RAF text 'RM', introduced to represent a separate engineering field, was the longest at 2921 words. Throughout this study each text will be referred to using its acronym in Table 1.2. In general, the texts classify their subject in a manner typical of technical text, providing readers with a taxonomic description (Martin, 1989; Halliday and Martin, 1993). The texts are not complete chapters, but they all constitute complete sections within a chapter. Each text was selected to represent the initial introduction to a topic, and so they tend to look forward (or Predict – Tadros, 1985; 1989) to following content rather than summarising previous points.

1.4 Outline of Remaining Chapters

The remaining four chapters are divided, roughly, into the textual systems of Participant, Theme and Information, followed by a concluding chapter. Chapters 2-4 follow the same outline. A review of the theory and research surrounding these concepts results in a model for textual analysis. The texts included in this study are subjected to the proposed model, and the results of the analysis are discussed, leading to further implications and conclusions.

Chapter two focuses on the two interdependent systems of Participant Identification and Participant Tracking. The highlight of the initial discussion, which covers a range of

approaches to reference, centres on an alteration to Martin's (1992) system of Participant Tracking where I allow Martin's taxonomy of semantic relations to substitute for Clark's notion of 'Bridging' reference. The text analyses support this decision. The results of the text analysis wrap up the chapter before considering how the current model might be improved.

Source	Title of Text	# words	Acronym
Black, B. J. (1997). *Workshop Practices and Materials*. London: Butterworth-Heinemann. p.183-7	Milling Machine	1020	MM
Buchla, D. & McLachlan, W. (1992). *Applied Electronic Instrumentation and Measurement*. Englewood Cliffs: Prentice Hall. P.37-40	Error, Accuracy and Precision	1909	EAP
Coope, S., Cowley, J. & Willis, N. (2002). *Computer Systems: Architecture, Networks and Communications.* London: McGraw Hill. p.252	Broadcast Networks	467	BN
Coulouris, G., Dollimore, J. & Kindberg, T. (2001). *Distributed Systems: Concepts and Design. 3rd Edition.* Harlow: Pearson. p.31	Architecture	486	Arc
Coulouris, G., Dollimore, J. & Kindberg, T. (2001). *Distributed Systems: Concepts and Design. 3rd Edition.* Harlow: Pearson. p.252-3	Security	710	Sec
Horowitz, P. & Hill, W. (1989). *The Art of Electronics*. Cambridge: Cambridge University Press. p.428,430	Amplifier Noise	385	AN
RAF Training Manual ref: RAF PTC CN 3787 1-1-6 06-528a/01/B50 1-1-7	Retardation Methods	2921	RM
Rappaport, T.S. (2002). *Wireless Communications – Principles and Practice*. Upper Saddle River: Prentice Hall. P.256-257	Frequency Modulation vs. Amplitude Modulation	1320	FMAM
Tannenbaum, A.S. (1995). *Distributed Operating Systems*. Upper Saddle River: Prentice Hall. P.246-248	The File Service Interface	871	FSI
	Total	**10,089**	

Table 1.2 Corpus of texts for analysis

The discussion in Chapter three is almost exclusively centred on SFL studies, as the topic of this chapter – Theme – has been researched more extensively in SFL than the other two topics. Consequently the discussion aims to make my stand clear on a number of controversial issues in Theme. This is followed by analysis of the same texts. The results of the analyses are discussed first in their own terms, and then in combination with the participant analyses from the previous chapter. A brief discussion of some shortcomings of the Theme-Rheme model are discussed.

The discussion of information structure in chapter four extends beyond SFL, and compares a range of definitions and approaches. The same texts are analysed and the three systems of text are combined to demonstrate how the proposed model operates. Chapter five then investigates this model further by attempting to explain how the model of information structure in written English differs from that in spoken English by examining the relevant functional, historical and psychological factors. I believe that the model will have considerable implications for linguistics in general.

Chapter 2 Participant Identification and Tracking

2.0 Participants, Theme and Information Structure

This chapter, and the following two chapters, aim to distinguish the information structuring roles of Participant Identification and Tracking, Theme and Rheme and Given and New information, following the division of labour in the textual metafunction proposed by Fries (2000; 2002) and depending heavily on Martin's (1992) discourse semantics. Initially, the textual systems of Participant Identification and Participant Tracking, which includes Presenting and Presuming reference, will be investigated, particularly as what is referred to as information structure in some theories can be accounted for by the functions of Participant Tracking and Reference. The model produced by this discussion is then applied to a selection of texts, and Participant Identification and Tracking are compared, with an emphasis on distinguishing their roles, functions and realisations.

Frequently the term referent is used to identify elements in text that can be carried forward and used repeatedly to enable discourse to progress based on what has gone before. The term referent, however, has come to imply a model of discourse dependent on a representationist view of knowledge. Section 1.2.4 outlined why a representationist approach to language should be avoided. Representationist views of language structure can only describe static relations between concepts which embody a meaning independent of language. This study uses the discourse semantic notion of Participant to enable a construing role for anaphora and related phenomena, rather than a static representation of a pre-determined system of meanings.

2.1 Participants in Text and Cohesion

A Participant is a major constituent in the Experiential metafunction – that part of language structure which represents ideas, actions and objects. Within the Experiential metafunction, a major clause consists of an obligatory process, recognised in traditional grammar as the verb, and usually at least one participant. A participant "can be defined as a person, place or thing, abstract or concrete, capable of functioning as agent or medium in TRANSITIVITY" (Martin, 1992, p.129), is any element that can function as (obligatory) Medium or (optional) Agent in a process in a clause, and is typically realised by a nominal

group (Martin, 1992). Not all nominal groups, however, are participants: introductory 'it' for the weather or adjectival phrases such as "*It* is unfortunate that he lost the game", negative groups as in "There was *nobody* there" and phrases in idioms (there is no *flash* and no *pan* being referred to when using the idiom '*a flash* in *the pan*') are a few examples where nominal groups do not coincide with participants in discourse (Martin, 1992).

Agent and medium are derived from an ergative analysis of the clause. A participant is defined by its role in the system of transitivity (more specifically, of ergativity), not by phoricity or reference; a participant is identified by analysing the potential ergative roles of the nominal groups in a clause. In conjunction with a transitivity analysis (which examines whether a process extends over more than one participant), a clause can also be analysed for ergativity, which examines the source or cause of the process (Halliday and Matthiessen, 2004). The Medium of the process is so named "since it is the entity through the medium of which the process comes into existence" (Halliday and Matthiessen, 2004 p.291) and is an obligatory function in most cases. The Agent is the most common optional function and is associated with the external cause of a process. In an ergative analysis, Beneficiary and Range (as in *the sea*, below) are potential participant roles (Halliday and Matthiessen, 2004 p.291):

> there may also be constructs like English prepositional phrases, the function of which is to bring in other potential participants but to bring them in indirectly, like *the sea* in [birds are flying] *across the sea.* (Halliday and Matthiessen, 1999, p.513)

That is, nominal groups in prepositional phrases do not have a role in ergativity, but they have the potential to be 'extracted' from the prepositional phrase and realise an ergative role. The second row in Fig. 2.1 illustrates which elements will be analysed: the focus in this study will be on potential and actualised participants.

Potential participants may also be realised by nominal groups derived from Circumstances or Processes. Nominal groups contained within a Circumstance (typically a prepositional phrase or, occasionally, adverbial group) will be analysed. Processes are also potential participants since they can be nominalised and become participants through grammatical metaphor (Halliday and Matthiessen, 2004). Processes and non-nominal Circumstances (layer 3 in Fig. 2.1) will be analysed only after they have been nominalised in the text (layer 3 in Fig. 2.1) so that they can perform the role of Potential Participant (layer 2 in Fig. 2.1). A non-nominal Circumstance will typically be an adverb, which can then be tracked

through a nominalisation. In this example, *repeatedly* is a Circumstance in the first clause, but is related semantically to *repetition* in the second.

> This is why she may ask you to read the same story or sing the same song repeatedly. Repetition also teaches her about actions and consequences and how one affects the other. (Available: http://www.childwellbeing.org/AtAGlance_pages/AAG-routine.html on 16 Oct. 2009)

Repetition here is analysed as a participant, and can be tracked back to *repeatedly*, although this is not a participant. Thus, a participant can be animate or inanimate, carry out the process or be performed upon, it can be seen or referred to in the real world or be a virtual entity construed only within the text, or it can be a process within the text or a part of the text itself.

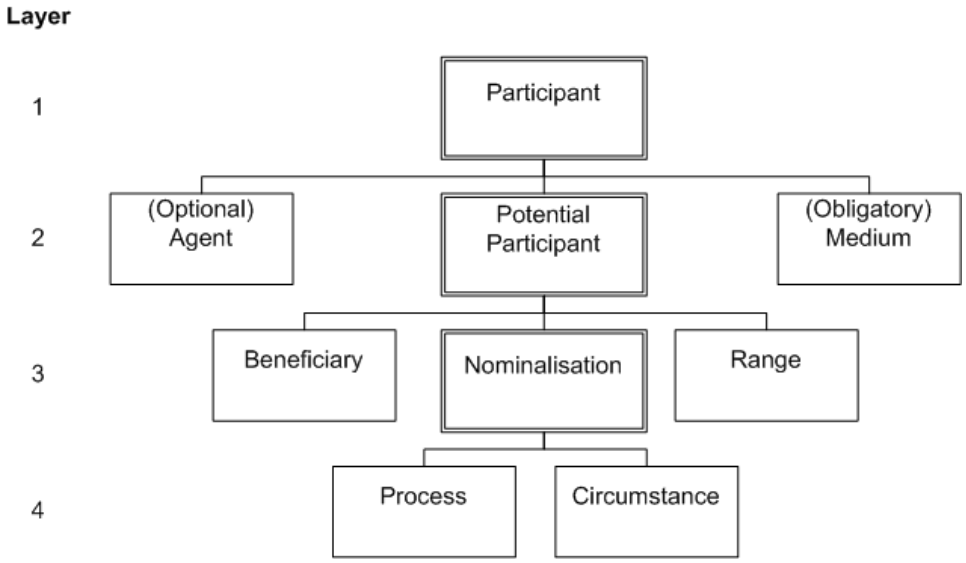


Fig. 2.1 Potential Participants and Participants in Text

Participants play an essential role in creating cohesive text. As participants are woven back into subsequent sentences in the text at various points they contribute to the cohesion of the text; participants need to be repeated in the text to create 'cohesive ties'. Halliday and Hasan (1976) consider only those ties that occur between sentences as cohesive, but semantic relations operate between participants both within and between sentences. As participants progress through various processes in the text they create 'cohesive chains', but cohesive chains and ties are not sufficient to create fully cohesive text (Halliday and Hasan, 1985). Participants need to interact with one another, and to interact with the various processes in the text in order to build interaction between the cohesive chains, producing 'cohesive harmony' (Halliday and Hasan, 1985; Hasan, 1984; *cf.* Hoey, 1991a): "variation in coherence in a text is the function of variation in the cohesive harmony of a

text" (Halliday and Hasan, 1985, p.94). One method of analysing the contribution of participants to the cohesion of discourse is through Martin's (1992) participant tracking. Each participant needs to be identified (see section 2.4.1) and analysed for phoricity (section 2.4.2) in order for participant tracking to occur. The following section reviews some explanations of how we manage to keep track of participants through discourse.

2.2 Research into Participants in Linguistics

Although there has been a great deal of research into linguistic phenomena such as anaphora and cohesion, which are related to participant tracking, very few studies focus on the identification and tracking of participants in text. Cohesive devices, including anaphora, are certainly related to participants, and are frequently used to identify and track participants, but few studies of cohesion directly employ participant tracking. One major contributor to the study of participants, Chafe, will be reviewed in more depth in chapter four, as many of his contributions relate more directly to Information structure. This section reviews the most frequently cited studies of relevance to participant tracking, focussing initially on the concept of 'bridging' and associated studies, before proceeding to research in computational linguistics. The ultimate aim of this review is to identify compatible concepts and models from other schools of linguistics that can be integrated into a systemic-functional analysis.

2.2.1 The Psychology of Participant Identification

Probably the most influential psycholinguistic model of text processing that reflects the tracking of participants through a text is the 'Given-New Contract' presented in full by Clark and Haviland (1977). The model proposes a Gricean *Maxim of Antecedence* and depends on the inference-based notion of *bridging*. These pragmatic concepts were based on a psychological approach to text processing, and are outlined below. This discussion of Clark and Haviland is continued in section 2.2.3 with studies that have been most influenced by their findings.

2.2.2 The Given-New Contract and Bridging

In a highly cited paper, Clark and Haviland (1977) base their argument for a *Given-New Contract*, and consequently for the need for *bridging* in inferencing, on linguistic and psychological evidence. Despite its name, Clark and Haviland's 'Given-New' contract is

mostly related to participant tracking (see section 2.4.2). It specifically claims to explain the psychological and pragmatic processes involved in antecedence. The 'given-new contract' reflects Grice's (1975) maxims and is defined as the 'maxim of antecedence':

> Try to construct your utterance such that the listener has one and only one direct antecedent for any given information and that it is the intended antecedent (Clark and Haviland, 1977, p.4)

Although Clark and Haviland name this contract 'Given and New', the maxim makes clear that their focus is on anaphora; all examples and research that they cite focus on anaphoric pronouns. One of the innovations in their research is to propose the notion of *bridging*, whereby a listener is able to recognise that there must be an antecedent and can locate the most suitable candidate, even though this may require an inference to be made.

The following discussion of the Maxim of Antecedence is critical for a number of reasons. The most important reason is that the model may have outlived its usefulness and may now be hindering progress, especially in the computational fields, even though the model has helped a number of researchers to assimilate a model of information structure, particularly in more formal fields of linguistic inquiry. Another reason is that great progress has been made in the 30 years since this paper was published, and an evaluation in the light of new findings is necessary. It also appears that many researchers failed to evaluate the proposals in the original paper, especially the psychological studies, in relation to the evidence that is offered. Further, it is essential to understand the differences between theories using the terms Given and New information, and to clarify terminological differences so as to appropriately compare and integrate findings from different perspectives.

A discussion of the evidence offered by Clark and Haviland (1977) is divided below into four sections, detailing: the methodology of the 1974 experiments that support the 1977 paper; arguments and evidence from other sources; inconsistencies in the terminology and arguments within the paper; and final comments covering more general issues (see also Moore, 2008a).

2.2.2.1 Methodology

Clark and Haviland (1977) rely heavily on the psychology experiments in Haviland and Clark (1974) to present evidence for the 'Given-New contract' and the maxim of antecedence. This section examines the methodology and the conclusions of those experiments.

In the first experiment, Haviland and Clark (1974) presented 16 Stanford University students with 57 pairs of single-clause sentences in sequence on a screen, and timed how long it took students to press a button to show sentence comprehension. The second ('target') sentences in this experiment were the same in each group of subjects, but they either repeated one of the nouns from the first ('context') sentence with a definite article, or were related semantically with a different noun in the first phrase. The only example that they provide is:

> "(a) We got some beer out of the trunk" and
>
> "(b) We checked the picnic supplies"

as 'context sentences', both of which were followed by

> "The beer was warm"

as a 'target sentence'. Clark and Haviland (1977) conclude with sequence (b) that "there is no direct antecedent, and so the listener must build a bridge" (p.21). That is, the listener (or, in the case of all their experiments, the reader) must make an inference between the definite marker and something in the context or co-text so that the definite marker has an antecedent. The two subsequent experiments follow a similar design.

There seem to be a number of flaws in the design of this and later experiments, some of which apply to similar experiments, but some of which relate specifically to this study. The first is the assumption that 17 Stanford University students represent the whole of mankind. There is evidently something different about such people, but how that might interfere with the results of this experiment is unclear until a more representative sample has been tested in the same way.

Being exposed to a large number of pairs of sentences devoid of context will surely produce its own effects, none of which are likely to reflect real language processing since we are rarely expected to make sense of language presented to us in this way. Subjects are likely to recognise patterns in the data after this many repetitions, and may even double-guess the tester, choosing either to support or subvert the experiment. Without any kind of check on understanding, apart from 'Do you understand the sentence', there is no way of knowing how accurate comprehension really was. This is particularly important in later experiments where many of the sentence pairs do not make sense.

The conclusions reached by Haviland and Clark (1974) need to be tempered by the method of their experiments. Because of the method of subjects pressing a button when they feel they understand,

> there may have been some instances of "incomplete" contextual comprehension, and so the absolute differences between the conditions cannot be taken very seriously. (p.519)

However, they do not take their own warnings into account when they conjecture:

> he searches memory for a matching antecedent to this Given information, and on finding it, attaches the New information to the antecedent. If he cannot find a matching antecedent, then he must (a) build some sort of bridging structure, (b) treat all information in the sentence as new and begin construction of a separate structure, or (c) attempt to recompute what is Given and what is New in the sentence. (p.518)

Clark and Haviland (1977) take these results very seriously when they also try to account for other research – research that could equally be used to justify another theory, *e.g.* schemata, or frames – as they use the research to justify the Given-New contract and the notion of bridging.

The fact that each sequence contains just two sentences that bear no relation to the students' own lives or any previous context calls into doubt whether or not the experiment accurately replicates real instances of the text comprehension process, inferencing, or bridging, if they occur. It seems incongruous to investigate a phenomenon that is so dependent on context by removing all context and co-text. Context (of culture, situation, and co-text) is used in normal communication to disambiguate potential misunderstandings. It is extremely rare for a reader to actually look back for an antecedent, certainly physically and probably mentally (Just, Carpenter and Woolley, 1980). It is quite possible, in fact, that a reader scans forward for anaphoric devices, and in the current co-text looks for participants that can be 'picked up' in the text, thereby resolving anaphora before they are encountered in a sentence. This is certainly what centering theories imply (see section 2.7.1), and would match reading processes more accurately since there appears to be very little looking back during reading while scanning forward in the text is commonplace (Rayner, 1998; Shebilske and Reid, 1979; Smith, 1985).

In the circumstances described by this experiment, it is unsurprising that subjects seem to look back to the previous sentences to resolve anaphora – they have no choice because no other context or co-text has been provided, apart from the setting of a psychology laboratory. It is in such circumstances that serious doubt has been cast on the methodology of the psychological laboratory (Jones and Elcock, 2001; Richards, 1996).

The methodological design, not the phenomenon under investigation, forces one result and there are no attempts to falsify the hypothesis.

2.2.2.2 Appropriation of other Arguments

As well as their own research, Clark and Haviland (1977) appropriate linguistic and psychological research carried out by others to support the 'maxim of antecedence', citing Chafe, Chomsky and Halliday to establish the syntactic reality of the categories Given and New. However, none of the definitions offered refer to the same linguistic units, or to the same functions in language. Clark and Haviland (1977) conflate Information structure with presupposition (and other phenomena), revealing a narrow interpretation of Halliday's definition of Given and New (see section 4.1 for further discussion). Their definition draws on pragmatic presupposition and implicature rather than on linguistic features, such as intonation (Halliday, 1967a) or reference. Clark and Haviland (1977) use arguments from linguistics to suggest that communication breaks down when the Given-New contract is violated by appropriateness, uniqueness or computability.

The following sequence is offered as an example of 'unacceptability' within the Given-New contract: "Two men were watching the dog. The one watching it laughed out loud." (Clark and Haviland, 1977, p.16) since it is not possible to compute a *unique* antecedent. This is not an example of "unacceptable" ordering of given and new information. It is contradiction. To be told in one sentence that two men were doing something and then that only one was doing it is unconnected to any 'Given-New contract' and so cannot be taken as evidence of its existence. Clark and Haviland argue that because the second mention of the man is not made unique it is unacceptable, and propose that by identifying the man in the second sentence as "the tall one watching it" solves the problem of uniqueness. Similarly, 'the tall man watching it', 'the tall one' or 'the tall man' would also make this unique, but would also mean that the second sentence no longer contradicts the first – an entirely different argument.

Clark and Haviland (1977) next offer a study of personal pronouns as evidence for the Given-New contract and the concept of bridging. Subjects were presented with hypotactic reporting clauses that included combinations of male and female and reflexive and non-reflexive pronouns followed by a 'checking sentence' that subjects were asked to indicate as being true or false. The research showed clearly that when pronouns were marked for gender and reflexivity, and when pronouns were physically closer to their antecedents,

processing of the sentence was faster. Clark and Haviland's claim that "these results further confirm the notion that the search for antecedents of non-reflexive pronouns begins in the same clause and goes backwards" (p.29), however, is still conjecture. There is no evidence for movement, backwards or otherwise. When comparing these claims against those offered by the *Centering* approach (see section 2.7.1 and 2.7.2), it seems quite possible to suggest that for each potential participant type that can be pronominalised, at any point in the discourse, the one that is most salient is the most-recently named participant, thus allowing faster processing to take place. Clark and Haviland concede that the 'backward search' model leaves a number of important questions to be answered.

2.2.2.3 Breaching and Violating the Maxim

Just as Grice's (1975) Maxims can be breached or violated, Clark and Haviland (1977) describe how the 'Given-New strategy' can be violated negligently, covertly or overtly. 'Negligent' violations, which result in a breakdown in communication, slow down the listener (or reader) because of an inability to identify the correct antecedent. As an example, an experiment (Bransford and Johnson, 1973 quoted in Haviland and Clark, 1977) using a paragraph with many hyponymic, or anaphoric, lexical items (e.g. 'things', 'groups', 'facilities', 'procedure') showed how such a paragraph is easier to understand when the 'empty' lexical items are filled with a topic that covers all of the passage. It would be equally easy to argue that this experiment supports the hypothesis for schemata, or frames (Emmott, 1994; Rumelhart, 1980; Tannen, 1979), but Clark and Haviland have only one explanation as to why the topic makes it understandable: "Apparently, it is the topic that enables the listener to compute the intended antecedents of each sentence in the paragraph." (p.33). There is no evidence that the intended antecedents in all of the sentences have been computed, nor that this is necessary for the subjects to claim that they have understood the paragraph.

Covert violations are distinguished by intent: they "are meant to deceive" (p.34). However, distinguishing 'negligent' from 'covert' violation by volition creates unfalsifiable categories. Lawyers and psychologists may spend time and effort deciding how to say something to intentionally mislead someone, but it is rarely easy to judge whether someone "is violating the contract and does not want the listener to know" or "simply misjudges what the listener does or does not know" (p.32). Either can be claimed, regardless of intention.

Finally, Clark and Haviland offer examples of overt (or explicit) violations of the Given-New contract. The violation is overt where the reader is 'dropped into' an ongoing narrative, "since the reader and writer are both aware that the reader cannot really compute the intended antecedents" (p.37). Clark and Haviland do not anticipate a breakdown in communication in this case, reasoning that the reader accepts the situation of limited ambiguity. They point out that as eavesdroppers we are in a similar situation, "we must be content with setting up antecedents by the Step 2 detour we called addition and hope that their true nature will become clear later." (p.37) In other words, even though an item is marked as having an antecedent, the reader or listener treats it as New. What makes these cases different from any other where Clark and Haviland proscribe ambiguity is not made clear. Earlier they assert that "On encountering … a pronoun the listener must compute its intended antecedent" (p.27). Yet, here they admit that this is not obligatory. When searching to resolve anaphoric devices that have no antecedent, Clark and Haviland claim that a reader asks the identity of participants: "The questions produce suspense and an impression of impending action" (p.37). It is not imperative, therefore, that every antecedent is 'computed'. In fact, it is quite possible that we frequently 'suspend' the solving of many anaphors. That we can suspend anaphoric resolution means that the demand to search for an antecedent may frequently be suspended; the maxim of antecedence is easily suspended and no bridge is built. Perhaps this is a question of register or context, and is acceptable only in the type of novels exemplified. Perhaps not. Clark and Haviland provide no guidance on these questions.

Significantly, if we are to presume that these violations are valid, then we are forced to re-examine all of the results for the experiments where experimenters are blatantly violating the Given-New contract, and the subjects are acting accordingly – knowing that these maxims may be violated. Under these circumstances, the only measurement used to justify the existence of bridging, (the length of time required to indicate comprehension), cannot be judged a reliable indicator.

2.2.2.4 Internal Inconsistencies

Looking in detail at Clark and Haviland (1977) there appear to be points where the arguments proposed and the analyses offered disagree with each other.

Haviland and Clark's (1974) second experiment tried to account for the possible effect of lexical repetition in experiment 1 by presenting new sequences to 10 Stanford students.

Subjects were presented either with the context sentence containing the same 'direct antecedent' (sentence a) as in the previous experiment or with a sentence similar to the following: "Andrew was especially fond of beer" followed by the target sentence "The beer was warm" (as in the previous experiment). Looking carefully at the results, we can see that the second sequence took longer to 'understand' than any in the previous (or following) experiments. According to Haviland and Clark "mere repetition of the critical noun is not enough to account for the results of experiment 1." (1974 p.516). However, offering a definite article for a Specific 'beer' in the second sentence when only Generic 'beer' is mentioned in the first sentence caused more confusion than a semantically related item (such as 'picnic supplies') (see section 2.4.1.2 on Generic and Specific). Clark and Haviland (1977) claim that "The indirect antecedent took about 140 msec. longer", but there was in fact no antecedent, since lexical repetition does not guarantee co-reference. In the case of these experiments, it is clear that Clark and Haviland (1977) do not distinguish between different types of reference, comparison and lexical cohesion (Halliday and Hasan, 1976) such as co-reference (identical referents), co-classification (distinct members of the same class) and co-extension (distinct members associated by semantic relations such as meronymy and hyponymy) (Halliday and Hasan, 1985). Had they done so, they may have found that these categories were the dependent variables in their studies.

This experiment, therefore, does not provide the evidence claimed by Haviland and Clark, since it should still be quicker to relate the repeated lexical item (beer – the beer) to the previous one than to relate it to a synonym, hyponym or similarly related item (picnic supplies – beer), if that item is in fact the same referent. The 'extra inferential step', (Clark and Haviland, 1974 p.516) that is conjectured to connect the two lexical items, may not even take place. Instead the subjects in the laboratory may just realise that, despite the lexical repetition, there is no connection between the first and second mention of beer, and indicate that they understand the sentences as such. In such instances it seems more appropriate to offer the option of 'I understand that this cannot make sense.' Without a reliable method of checking comprehension it is impossible to say.

The timings provided by subjects for comprehension of sentence pairs reveal a minimum of 1000ms for direct and indirect conditions and the time to make an indirect reference is only 7% longer than for direct reference. It is possible therefore that a 'bridge' is also made for direct antecedents. I would suggest that the discrepancies in Haviland and Clark's (1974) data are a result of ignoring the semantic features of what they class as

'bridging' or inferencing. It would appear that complete (or direct) repetition is the easiest relationship to recognise, followed by inflexion, derivation, different types of co-hyponymy (including types of contrast), hyponymy and hyperonymy. These and other semantic relations are represented in Fig. 2.11, and discussed in section 2.4.2.4. A semantic scheme could explain the apparent discrepancies in the data, and would account for some of the 'timelag' features identified in their experiments. It is possible therefore that a 'bridge' is made between complete, or direct, repetition of lexical items or pronouns in exactly the same manner as in the other types of matching participants in text, in which case there is no essential difference between 'direct' and 'indirect' reference, and possibly no need for the category of 'bridging'. That is, the difference between matching a referent with its 'intended antecedent' may be different in degree but not in kind between a pronoun, a repeated or derived lexical item, a synonym or hyponym, or an item related by meronymy or by collocation. This would require a major re-evaluation of the 'Given-New contract'.

2.2.2.5 Final Comments

The discussion above has examined some of the psychological and linguistic evidence presented by Haviland and Clark (1977). While this is not the first criticism of Clark's approach (see Asher and Lascarides, 1998; Matsui, 1999; Wilson and Matsui, 1998), I believe it is the first to question the distinction between direct and indirect reference.

I have attempted above to outline why the data provided by the psychological experiments are insufficient to posit a 'bridging' or inferencing process. The data show a clear difference between the average time taken for subjects to press the button for sentences with a repeated lexical item than for those with a semantically-related item. The interpretations placed on that fact, however, are almost limitless. The fault here appears to be in the logic applied to the experimental results. While it may be true that

> if there is no suitable antecedent for the Given information, the listener should be very slow to claim he comprehends the sentence (Haviland and Clark, 1974 p.514)

it does not logically follow that any time lag in comprehension must be the result of searching for an antecedent. It is clear that some of the sentences took longer for students to claim comprehension, but since the experimental design cannot disprove the null hypothesis the data cannot claim to prove the existence of inferring a bridging relationship. Differences between experimental sentences could easily be explained with a more comprehensive model of reference, allowing for distinctions between co-

classification, co-extension and co-reference (Halliday and Hasan, 1976; see section 2.2.2.5).

One argument here is that there is no Given-New strategy, and that 'bridging' does not occur. Certainly, I am proposing that the evidence provided in Clark and Haviland (1977; Haviland and Clark, 1974) does not prove its existence, mainly as a consequence of inappropriate evidence, a refusal to consider alternative explanations, results that reveal a negligible difference between conditions, and poor experimental design. These would not be the first experiments into linguistic phenomena that were questioned (see the 'Wason' debate by Sperber *et al.*, 1995; Fiddick *et al.*, 2000; Girotto *et al.*, 2001; Sperber and Girotto, 2002), and there is no reason why experimental psychology should not become more reflexive and critical, particularly in its methods (Jones and Elcock, 2001).

Finally, if we are to accept that 'bridging' does occur, it is not clear how this will help us to analyse discourse. Since Clark and Haviland (in 1977 or thereafter) provide no details as to how the bridging process operates, its features, criteria for successful bridging, its location in a schemata or greater theory of cognition (except to characterise it as a form of inference), we are led into a theoretical dead-end. Offering a psychological basis to the concept of bridging consigns the issue to an impenetrable 'black box' of cognition.

To summarise, Clark's work (Clark and Haviland, 1977; Clark, 1977) makes constant reference to Haviland and Clark (1974) to justify the concept of bridging. However, as I have attempted to demonstrate, the evidence offered for this phenomenon falls far short of establishing its validity. Many other possibilities could explain the poorly designed experiments in Haviland and Clark (1974), and the concept itself stands on theoretical ground that has since been criticised, even within pragmatics (Wilson and Sperber, 1986; Giora, 2002). If we are to adopt a psychological explanation for anaphoric reference, we would be advised to look beyond one based on *bridging*.

2.2.3 Consequences of Haviland and Clark's Work: Given-New Taxonomy

One of the studies that borrowed directly from Clark and Haviland's work, using the unanalysed concepts of bridging and the 'Given-New contract' was Prince (1981), which offers a taxonomy of Given and New (see Fig. 2.2 below) in place of the 'confusion' surrounding these terms introduced by Halliday (1967a), Chafe (1970; 1976), and Clark

and Haviland (1977). Although these terms are derived from the very different linguistic perspectives, Prince offers no justification for organising them in this way. The taxonomy is designed to cover *bridging* inference (Inferrable), and Indefinite (New) and Definite (Evoked) reference.

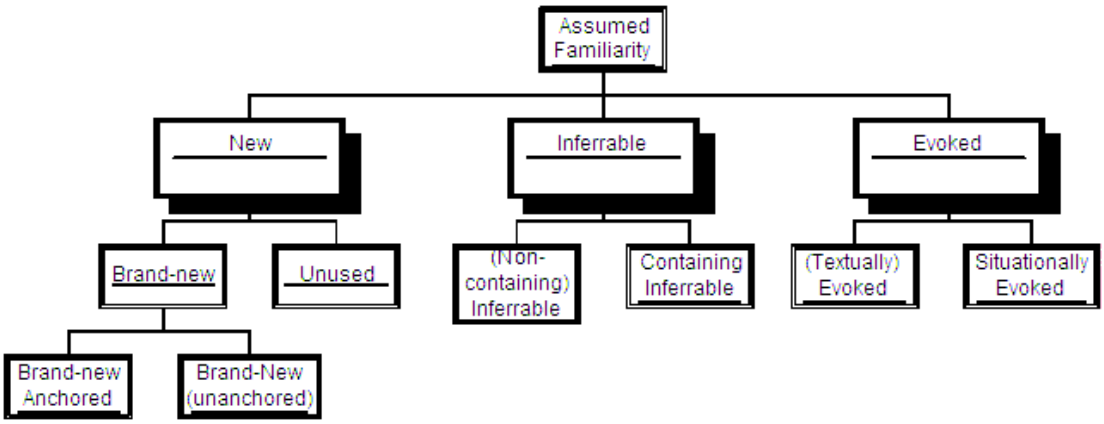


Fig. 2.2 Taxonomy of Given and New (from Prince: 1981 p.237)

While the taxonomy certainly has an intuitive appeal, and Prince (1981) does apply the taxonomy to text, it is ultimately unsubstantiated, especially as Prince (1992) appears to undermine the model. This is significant because although many other studies have taken Prince's taxonomy and based further work on it, none have attempted to demonstrate its psychological or linguistic validity. If the arguments supporting Prince's work can be shown to be fundamentally flawed, then the taxonomy and attendant studies must also be questioned. In a study of classroom language, for instance, Riesco-Bernier (2002; Riesco-Bernier and Romero-Trillo, 2008) attempted to apply Prince's taxonomy to a large corpus of spoken language, but found that categories based on more functional criteria explained the data far more adequately.

Prince (1981) dismisses the term 'shared knowledge', claiming that accessing 'shared knowledge' in a text is evidently unfeasible and "is not what ordinary, non-clairvoyant humans do when they interact verbally" (1981 p.232) because a speaker (or, even worse, writer) does not possess the omniscience required to be aware of the listener's (or readers') belief set and state of knowledge. Prince prefers the term 'Assumed Familiarity', but later (Prince, 1985) also abandons this term. Following Prince's line of argument, speakers and writers may be unable to predict the degree to which someone will make presuppositions about, or infer, certain knowledge, or the type of knowledge that the listener/reader has in order to make presuppositions or inferences, and yet Prince is satisfied to leave the concept of 'inferencing' unexamined.

More serious, however, is Prince's dependence on Clark's view of perception, cognition and attention. Prince (1992) still depends on the model of attention proposed by Clark (Clark and Haviland, 1977; Clark and Marshall, 1981) to support the proposed categories. It is the Clark studies that provide Prince with the categories of 'hearer-old' and 'hearer-new' in Prince (1992) and 'evoked', 'unused' and 'brand-new' in Prince (1981). As described above, these studies need verification before basing any further theories on them.

While Prince (1981) constantly refers to discourse and dialogue, all examples are invented and no longer than two sentences. In Prince (1992), however, a selection of the variables from the scheme above, namely discourse-old (Textually evoked), hearer-old (situationally evoked) and inferrable, are examined in terms of subject-hood in a text. In the study Prince identified just one factor that was consistently related to subject: discourse-old. That is, hearer-old is not an independent variable in terms of subject-hood in the text examined. Although this is a very small sample of language, and only one aspect of it is examined, Prince's own studies call into question some of the categories proposed in the 1981 paper. For that reason, it is essential that studies re-examine their dependence on Prince (1981).

2.2.4 Consequences of Haviland and Clark's Work: The Givenness Hierarchy

Gundel, Hedberg and Zacharski (1993) devised a Givenness hierarchy, whose "main premise is that different determiners and pronominal forms conventionally signal different cognitive statuses" (Gundel, Hedberg and Zacharski 1993 p.274), as illustrated in Table 2.1. Acknowledging the influence of Prince's (1981) Familiarity Scale, they point out that their hierarchy shows greater validity as each category, from left to right, subsumes the previous one:

> Since each of the cognitive statuses in the Givenness hierarchy entails all lower statuses, a particular form can often be replaced by forms which require a lower status. (Gundel, Hedberg and Zacharski 1993 p.294)

The realisations of each category, given below in table 2.2, are neither exclusive nor complete. *This* and *that*, for example, in each case are meant only as examples and Gundel *et al.* (1993) point out that the identical form of *this* can in fact have a range of 'functions' that may be realised by separate words in other languages.

in focus>	activated>	familiar>	uniquely identifiable>	referential>	type identifiable
it	*that* *this* *this* N	*that* N	*the* N	indefinite *this* N	*a* N

Table 2.1 Givenness Hierarchy (Gundel, Hedberg & Zacharski 1993 p.275)

While the model has been tested to a limited extent (*e.g.* Gundel, Hegarty and Borthen 2003), more analysis is advised especially as Gundel, Hedberg and Zacharski (1993) insist on a representationist model. Typical examples include the definitions of the categories in the Givenness Hierarchy:

> **FAMILIAR**: The addressee is able to uniquely identify the intended referent because he already *has a representation of it in memory* (in long-term memory if it has not been recently mentioned or perceived, or in short-term memory if it has) …
> **ACTIVATED**: The referent is represented in current short-term memory. *Activated representations may have been retrieved from long-term memory*, or they may arise from the immediate linguistic or extralinguistic context. (Gundel, Hedberg and Zacharski, 1993, p.278, emphasis added)

These quotations combine a representationist model of reference theory with phoricity to produce definitions that seem very similar. Both categories depend on a representation being present in either short- or long-term memory, so the distinction between these categories seems difficult to verify empirically. (See section 1.2.4 on the tension between a representationist model of language and a constructivist model of reference).

The Givenness Hierarchy demonstrates a system for reference where a reflex in meaning, expressed here in psycholinguistic terms, is found in the grammar, and so offers some promise for an SFL approach. Some caution is required before adopting this model, however. The Givenness hierarchy depends on Prince's (1981) work, which has not been demonstrated to be empirically accurate, and which in turn is based on the problematic notion of bridging (see section 2.2.2). It may be possible to apply the two SFL systems of Participant Identity and Participant Tracking to the Givenness Hierarchy proposed by Gundel *at al.* to classify Given and New in the sense offered by Chafe (e.g. 1970) in order to reduce confusion with the categories essential to this study. More importantly, though, a similar but more comprehensive classification of determiners is provided by Martin (1992) and discussed below (section 2.4).

This review of studies that were intended to provide a psychological account of participants in text has revealed how errors in seminal papers have been left unexamined and incorporated into later theories, leaving them vulnerable to critique. Within formal and pragmatic schools of linguistics, an attempt to establish a psychological basis for participant tracking, for Clark's (1977) view of Given and New, has to my mind been fairly unconvincing because of an insistence on representationist models of language and unsubstantiated theories of the mind. Clark *inter alia* and Prince have so far been unable to establish a credible role for the attention state of the hearer within a Gricean pragmatic paradigm. We turn now to approaches that use computational methods to identify any useful findings for the study of participant tracking.

2.3 Computing Participants

How we manage to identify the correct participant has been problematic for most formal linguistic theories (which often label a Participant a 'referent'). While the applicability of Halliday and Hasan's (1976) theory of cohesion was clear to many language educators with a background in formal linguistics (*e.g.* Stotsky, 1983; Witte and Faigley, 1981), it has not proved easy to integrate cohesive devices into a theory of linguistics that defines language as acceptable or unacceptable sentences, largely because the study of cohesion requires text – multiple sentences – in order to work. However, with the development of greater processing power in computers, there has been a surge of interest in formal approaches to anaphora – one method of tracking participants in text. This section outlines progress made in research into statistically-based computational methods of participant tracking, which has been mostly tested in the area of anaphora resolution.

Mitkov (2000; Mitkov *et al*, 2001) suggests that earlier studies that relied on "the representation and processing of various types of linguistic and domain knowledge" (2000, p.130) had failed because of an inability to accurately represent such knowledge computationally. He also criticises studies that require parsing of any kind as they "may render implementation more complicated and might slow down the searching process." (2000 p.130) He then reports on various studies using statistical approaches to anaphora resolution that he claims have proven increasingly effective in computationally resolving anaphoric pronoun reference.

In a seminal paper for approaches using the type of shallow processing described by Mitkov (2000), Lappin and Leass (1994) report an 86% resolution rate for third person

pronouns and lexical anaphors. Even with this narrow range of cohesive devices, the procedures described in Lappin and Leass (1994) are long and complex, after the corpus has initially been parsed using the Logic-Based Machine Translation System (McCord, 1989a; 1989b quoted in Lappin and Leass, 1994). The results of Lappin and Leass's (1994) study do not compare favourably with the more discourse- or pragmatic-based formal studies typified by *Centering* theories (see 2.8.1), but the 1994 study paved the way for other statistically-based studies.

Bos (2003) introduces the theory of Government and Binding into a theory of anaphora resolution by using a framework of presupposition. In a bottom-up approach to anaphora resolution, each lexical item is taken to represent part of the context within the sentence. As more context is discovered though co-text, the number of presuppositions reduces. Mitkov's (*e.g.* 2000) research follows in this tradition, relying heavily on tagged text. It also bears a striking resemblance to the Centering approaches, described in section 2.7.1, but does not report as high a rate of anaphora resolution as the Centering studies.

Despite all of the advances described above, before applying the findings of these computational results to linguistics, there are still a number of reservations covering issues of methodology and types of anaphora, the aims of the research, and theoretical background. The methodology of these studies is mainly computational. The advantage of this is that, through algorithms, researchers must make their method of anaphora resolution explicit. The disadvantage is that this seems to be possible only by taking a narrow view of language and of anaphora. In most of these studies, the highest rates of resolution occur when only third person pronouns are considered and resolved. This means that the majority of anaphoric devices, and an even larger set of cohesive devices used to track participants, are ignored in these studies.

One aim of research in computational linguistics is to use a computer to analyse text. This is likely to be most effective by combining an understanding of linguistics with the best approach available to a machine. The resulting model, however, is unlikely to mirror the approach that a human uses to process text because human cognition does not reflect a computational model (Edelman, 2004). Even when machines can analyse text efficiently, separate research will still be required to reveal the processes that are used by humans.

Perhaps the most disappointing research is represented by attempts to integrate more formal theories into a study of reference. For instance, when Bos (2003) asserts

"Presuppositions are an important linguistic device, because when conveyed in utterances, they put constraints on the discourse context" (p.180), he suggests that the linguistic realisation constrains the context. While linguistic realisations may constrain the co-text (or the discourse context), it is the context of situation and the context of culture that puts constraints on the utterance. These are precisely the contexts that are ignored in formal approaches to linguistics. When attempting to comprehend sentences in isolation of their context, it appears that "there are certain lexical items that give rise to presuppositions, whereas others do not" (Bos, 2003, p.180). The main question that then arises is how to identify those items that do not give rise to presuppositions and, for those that do, where to stop making presuppositions:

> Hence, in complex sentences there is *no systematic way* for dealing with presupposition triggers, as *sometimes* subparts of complex sentences carry presuppositions that are cancelled in the main sentence. (Bos, 2003, p.182, emphasis added).

This seeming lack of systematicity results in arbitrary analysis, and so anaphora resolution is no longer formally defined. A theory of language that requires the integration of context with language will not depend on a theory of presupposition to approach generation or processing (Halliday and Hasan, 1985; Sinclair, 2004).

In conclusion, the statistical-based approaches to anaphora resolution offer some interesting studies, some of which show a remarkably high level of success, and can only encourage all models to be as explicit as possible in order to implement schemes computationally for comparison. Success in computational anaphora resolution, however, normally comes at the price of ignoring a wide range of important cohesive devices and generally just concentrating on pronouns.

2.4 Participant Identification and Tracking

The aim of this section is to outline the systemic-functional view of Participants in text, looking first at Identification, and then Tracking. The entry condition for the systems below is a Participant in discourse. The description and definition of participants was provided in section 2.1, and a review of related theories was provided in sections 2.2-2.4. The major source of the following description derives from Martin (1992). The following section (2.5) will outline those aspects of other theories, reviewed in the preceding discussion, which are compatible with the systemic-functional model, as it is presented below, and which may modify the model proposed by Martin (1992).

2.4.1 Participant Identification

This section outlines the contributions made in *English Text* (Martin, 1992) to the study of reference in discourse. The model will then be evaluated against the corpus of texts for this study, to see if it can account for the choices made in the texts (section 2.6). Martin (1992) acknowledges the significant contributions made to reference theory by Halliday and Hasan (1976, 1985), but is keen to extend Halliday's (1985) grammar from the clause to the level of discourse. Thus, although many of the realisations that are examined are the same as Halliday (1985), the perspective taken in Martin (1992) is quite different. Using the ideational, logical, interpersonal and textual metafunctions as a guide, Martin approaches reference from the perspective of discourse semantics.

Each participant in discourse needs to be identified by distinguishing it from other participants. Failure to distinguish participants results in ambiguity. One of the major resources for identifying participants in discourse is through the use of reference (Martin, 1992). Participants may be identified with Presenting or Presuming reference. *Presenting* reference is used to *present* a participant to the text. Presenting reference is used to indicate that the speaker does not expect the identity of the participant to be recovered from the context. In contrast, Presuming reference is phoric. Martin (1992) uses the functional term Presuming because "Phoric items will thus be described as *Presuming* information from their context" (p.101, original emphasis) (see section 2.4.2 for further discussion of phoricity). Presuming reference indicates that a participant's identity can be tracked to the context of text, situation or culture. As part of their Identification, each participant may also indicate Generic or Specific, and Comparative reference. These three systems are simultaneous, and will be discussed below. First, however, the options that are available prior to these in the system of participant identification must be described.

2.4.1.1 Neutralised and Generalized Reference

It is possible to avoid the systems of reference and participant identification almost entirely, such as in the "Compressed English" (Sinclair, 1988) of certain discourse colonies (Hoey, 1986), including encyclopaedias (Bloor, 1983), many dictionaries, and newspaper headlines. In these instances, the identification system is Neutralised, and participants are generally not marked for phoricity, as in the heading for this section. Thus, while there may be cohesive relations, such as lexical ties, substitution and ellipsis,

the resources of reference are generally unavailable to identify and track participants. If the system is not Neutralised, it is Effected.

When a grammatical participant does not coincide with a semantic participant – when there is no participant to be identified – the full system for participant identification is not entered, and so the choice of Generalized reference comes into effect. Generalized reference is distinguished from Generic or Specific reference because it does not refer to a participant at all. Generalized reference is exemplified by the semantically-empty *it* in "It's snowing", or the use of *one* to refer to nobody in particular (Martin 1992 p.119). If the participant is not Generalized, it is Specified, and enters into the detailed system of identification. Fig. 2.3 shows options prior to the choices that allow Participant Identification. Participant Identification starts at the Specified point since it is here that an individual participant can be distinguished from others.

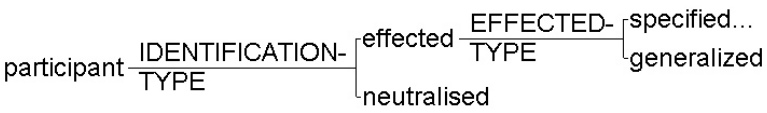


Fig. 2.3 Initial choices in Identification System

2.4.1.2 Specific and Generic

Specified is the entry point for the simultaneous choices of Specific-Generic (the focus of this section), Presenting-Presuming (section 2.4.1.4) and Comparative (see section 2.4.1.3 and Fig. 2.4, where "-t" represents no comparison choice selected). Although these three systems are simultaneous, there is one restriction on co-selection. Presuming reference is simultaneous with both Specific and Generic, whereas Presenting can only select Specific. Generic is not associated with Presenting reference, because the introduction of a participant using Generic reference necessarily presumes knowledge of that item – the knowledge originating in the context of culture rather than in the context of situation. Consequently, Generic reference can be realised by all forms of reference (*a lion*, *the lion* and *lions*), except a definite article with a plural noun (*the lions*) as this is the only option for a Specific plural.

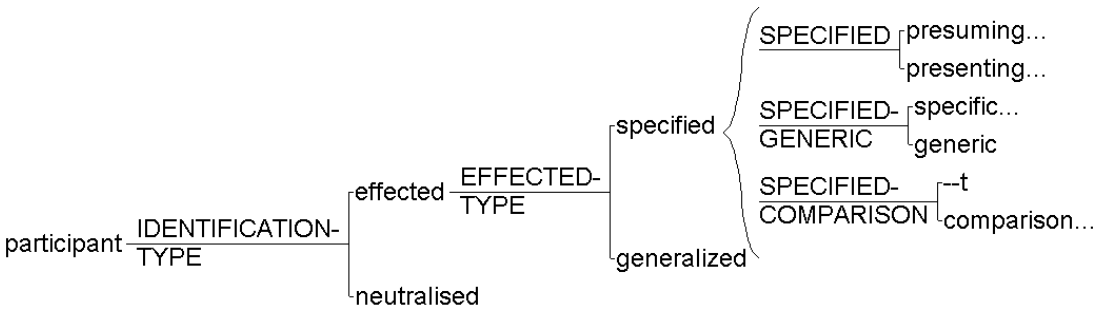


Fig. 2.4 Simultaneous Systems of Specified Participants

Generic reference often realises participants by using a general class of noun. Generic reference tends to be selected more frequently in the register of technical text than its normal low frequency (Martin, 1992 p.103). One test for Generic class of noun versus Generalized reference is that of pronominalisation; Generic nouns can use singular pronouns, classing them as a participant in the clause, while Generalized reference cannot. The term Specific is used for participants that can be distinguished from other participants in the context, and is the unmarked type. Generic and Specific reference can be distinguished in terms of the location of their referential value. Specific terms depend to some degree on the local context, and so are marked as such, while for participants marked with Generic reference "their context is in effect simply that of knowledge of the language being used." (Martin, 1992, p.103). Martin (1992) does not indicate any choice more delicate than Generic.

The distinction between Generic (or general knowledge) and Specific (to local context) may turn out hard to substantiate at times, especially across different registers. If we are to develop a model of language based on a constructivist model, we must resist models that encourage a 'representationist' view of knowledge and learning by developing active constructivist models. In a construing view, 'knowledge of the language' can be seen as a function of the frequency of language exposure and use (Bybee, 2006; Ellis, 2006; MacWhinney, 1998). That is, each idiolect is the product of the interaction between the learner's linguistic environment and their cognitive abilities. 'Knowledge of the language' (or languages, or language varieties) is thus dependent on individual experience. If 'Specific' means context-dependent, then the difference between knowledge of the language and reliance on the context will vary noticeably between individuals. For instance, in areas of specialist language, or rapid language change (such as in the area of new technologies (Bloor, 1998)), the amount of 'knowledge of the language', and the amount of 'Specific', i.e. context-dependent, knowledge will vary greatly. Experts will use

presuming reference for items that novices find alien. Novices – those not fully socialised into a discourse community – often depend on co-text for the items that experts consider part of the context of culture. In the following passage, an abstract from a review article, the Generic items are italicised, and Specific items are underlined:

> This review describes some recent, unexpected findings concerning *variation in spatial language across cultures*, and places them in the context of *the general anthropology of space* on the one hand, and *theories of spatial cognition in the cognitive sciences* on the other. There has been *much concern with the symbolism of space* in *anthropological writings*, but little on *concepts of space in practical activities*. (Levinson, 1996 p.353)

Generic reference is used here to locate an archetypal reader – one who shares a context of culture containing the Generic terms that are presumed, and is interested in the Specific items, being presented. Choice of phoricity may accordingly indicate solidarity for a particular discourse community and may also work to exclude others from that community, and from knowledge (Lassen, 2003).

If variation in level of knowledge results in different phoricity, then choice of phoricity becomes a semantic choice, as claimed by Martin (1992, p.102; Fries, 2000), and not a grammatical requirement irrespective of context. Generic reference is chosen if the speaker/writer makes the presumption that a participant does not need to be presented because it is evident in the context of culture. Consequently, all text is culture-specific, and the culture and ideology of the writer is revealed in choices such as the reference system. Thus, a model that attempts to explain the systems of reference and determiners without considering all context and co-text will be unable to account for these choices.

2.4.1.3 Comparison

If Comparison is selected, it can be realised in a variety of ways, including as a post-deictic (*last* in *the last house on the right*), numerative (a quantifying determiner like *fewer*) or epithet (a 'describing' adjective like *smaller*). Martin (1992, p.118) divides the system of Comparison into the two simultaneous choices, here designated by the functional terms Similarity and Specification (see Fig. 2.5 for the system and typical examples). Similarity divides into Difference and Semblance, which is divided into Identity and Similarity. Specification is classified into the choices of General and Experientialised, itself divided into two simultaneous systems. The first distinguishes the unmarked Experientialised from

Experientialised: Purposive, and the second distinguishes Comparison: Quality from Comparison: Quantity.

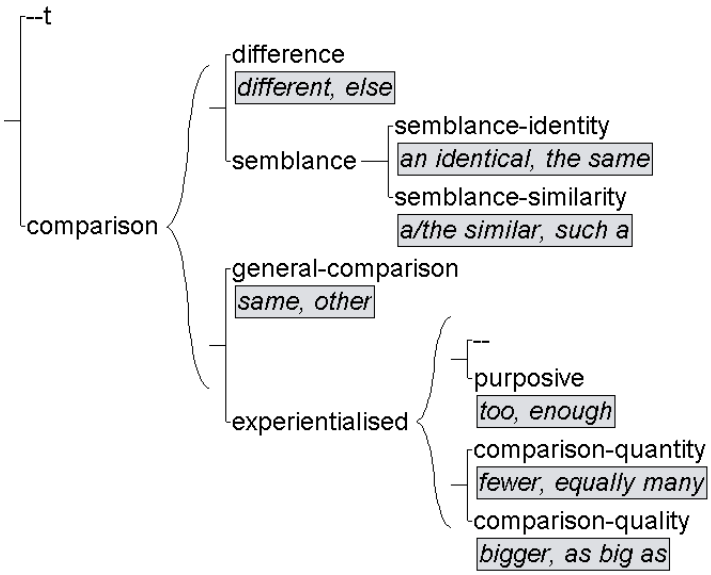


Fig. 2.5 Options and Typical Realisations in the Comparison Network

Comparative reference has remained problematic since Halliday and Hasan (1976) as it is often simultaneously presenting and presuming; there will always be a participant that is presumed as well as a participant (or a feature of a participant) that is being presented or presumed. It is, thus, distinct from choices in the presenting-presuming network. This explains why for Halliday and Hasan (1985) Comparative reference is both co-referential and co-classifying. When two items are compared, they are grouped into the same class for the purpose of comparison (co-classifying), and the presumed participant is re-invoked in the text (co-referential) while also possibly presenting a new participant. From a discourse semantics perspective, Martin (1992) deals with this dual purpose by placing Comparison in Relevance phoricity ("the identity of one or more participants related to the participant being realised is recoverable" 1992, p.100), which can also be realised in Superlatives, Conjunction and Continuity, as opposed to Reminding phoricity (typically realised by presuming reference) and Redundancy phoricity, typically realised by ellipsis, substitution and intonation.

The following passage exemplifies how comparative can be realised as numeratives:

> *Some forms of noise* are unavoidable (e.g. real fluctuations in the quantity being measured), and they can be overcome only with the techniques of signal averaging and bandwidth narrowing, which we will discuss in chapter 15. *Other forms of noise* (e.g. radiofrequency interference and "ground loops") can be reduced or eliminated by a

> variety of tricks, including filtering and careful attention to wiring configuration and parts location. *Finally*, there is noise that arises in the amplification itself, and it can be reduced by through the techniques of low-noise amplifier design. (Horowitz and Hill, 1989, emphasis added)

The use of *some* in 'Some forms', being an incomplete set, prospects to there potentially being other forms. This potential prospection is fulfilled with 'Other forms'. 'Other' does not mean anything on its own; it is defined in comparison with the original group. The use of 'Finally' explicitly completes the prospection, and is in comparison to the previous two groups, although it is not a participant. All of these comparative forms introduce sub-sets of a larger participant – 'forms of noise'. The presuming function of comparison is shown here as all belong to the same group of 'forms of noise', while the presenting function is shown by 'some', 'other' and 'finally' which introduce new sub-types.

2.4.1.4 Presenting and Presuming

The Presenting and Presuming options represent the major choices in participant Identification. While the Comparison sub-system need not be selected, and Specific is the unmarked (or 'default') choice compared with Generic reference, Presenting and Presuming choices are made constantly in all registers. The main choices within Presenting (Martin, 1992, p.104*ff*), with typical realisations, are illustrated in Fig 2.6, and discussed below, before the Presuming options (Fig. 2.7 below).

When a speaker/writer chooses Presenting reference for a participant, they indicate that they do not expect the audience to recover the identity of the participant from the context; the participant is being explicitly 'presented' at this point in the text. Although the choice of Presenting reference introduces a participant as new, independent of whether that participant has already been mentioned, Presenting reference can use pronouns. Both of the major divisions, of Total and Partial, are sub-divided into nominal and pronominal with realisations such as *everybody* and *something* for new participants that cover all of a group of people or an unspecified amount or number of objects, respectively.

Martin (1992) offers the hypothesis that the unmarked Presenting choice, resulting in *a* /ə/ or *some* /səm/ (i.e. realised with an unstressed 'schwa' vowel), contrasts with the stressed, marked Particular role accorded a participant being presented with *some* /sʌm/, as in "Some guy walks in and says…" in spoken narrative, or with a Major Role realised, for example, by *a certain* or *this* (*cf.* "This guy walks in and says…"). The order of the

marked choices of Unrestricted, Non-particular, Particular and Major Role (Fig 2.6) is hypothesised by Martin (1992, p.108) to form a cline of 'centrality' for presenting new partial, nominal participants. Each choice alerts the listener/reader to the increasing centrality of the new participant to the discourse – an Unrestricted Participant will be less central than a Participant presented as a Major Role.

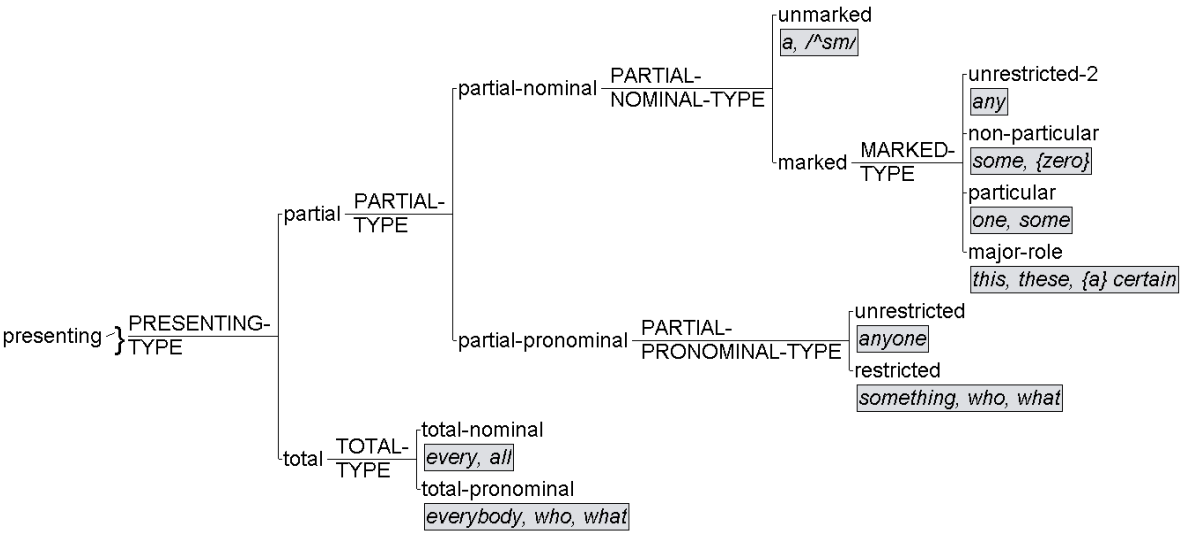


Fig. 2.6 Options and Typical Realisations within Presenting Reference (based on Martin 1992, p.104*ff*)

The first distinction in Presuming reference (Fig. 2.7) is between Unique and Variable. Unique reference means that only one is possible, and explains why most proper nouns are referred to without an article – the listener is expected to know which *John*, *London*, or *Microsoft* is being discussed without being told, because there is only one in the context. Note that if a proper noun is considered to be potentially not unique it can combine with an article, e.g. "the John that you met at the party" or "Do you mean *the* David Attenborough?" When a noun is unique it is signalled as such with the definite article (*e.g.* the sun), regardless of whether it is the first or fifth mention.

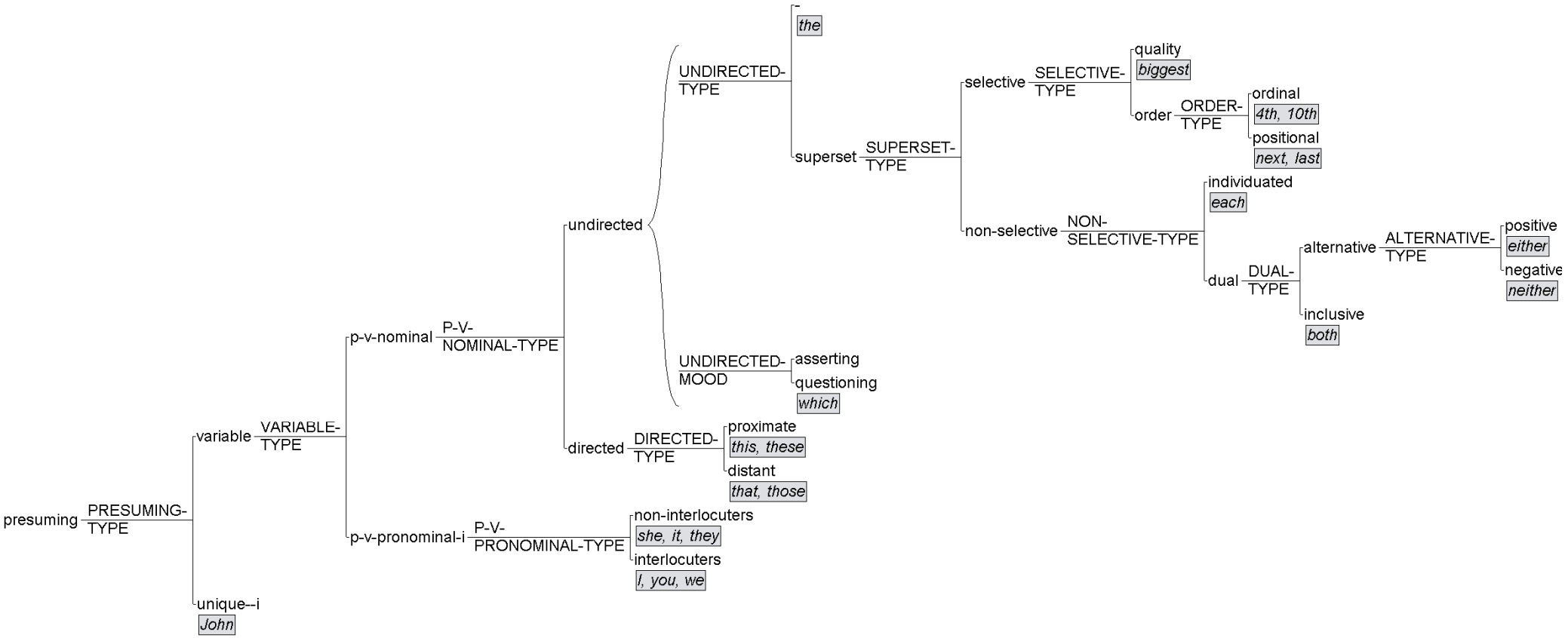


Fig. 2.7 Options and Typical Realisations within Presuming Reference

Presuming, Variable reference is subdivided, like Presenting, into nominal and pronominal. Pronominal realisations are divided into Interlocutors, such as *I*, *we* or *us*, and the Non-interlocutors *it*, *she*, *them* etc. Presuming, nominal reference can be Directed, which is then classified into Proximate, possibly realised by *this* + noun, or Distant (*that*). The unmarked choice for nominal Presuming reference is Undirected, and typically realised by *the* (which accounts for approximately 6% of all running text in English[1]). Within Undirected there are further choices available when a Superset is involved. The Superset choices divide into Non-selective and Selective, which is subdivided into Quality (*e.g. biggest*) and Order. Order is further classified into Ordinal (*1st*, *3rd*, *nth*) and positional (*e.g. next*, *last*). When a Superset is Non-selective, the choices divide into Individuated (*e.g. each*) and Dual. Dual then splits into Alternative and Inclusive (*e.g. both*). Alternative, itself, then divides into positive and negative, resulting in a network that displays a high degree of delicacy. These choices, and typical realisations, are charted in the network in Fig. 2.7. Appendices 2.3 and 2.4 display examples of choices from these networks, as well as Participant Tracking choices, from two texts in the corpus. The analysis is discussed in section 2.6.

2.4.2 Participant Tracking

Participant Identification may be seen as a preliminary to Participant Tracking (participants need first to be identified before being tracked), or as the static counterpart to the dynamic Tracking analysis; Participant Tracking is relative and depends on the developing context and co-text, while Participant Identification need not. As participants are identified and presumed through a text, it is possible to track each one as they interact with the process in each clause. Participant Tracking depends largely on the grammatical resources of phoricity, as well as lexical relations. Participant Tracking is a key feature in our understanding of the textual metafunction as it provides options for interactions with Theme and Information structure.

2.4.2.1 The Context of Culture – Homophora

The system of phoricity guides the reader to the location of the identity of a participant. If a participant can only be understood with reference to another entity, the question that arises is where the other entity can be found. In some cases, "participants can be treated

[1] In the results from the Bank of English provided with WinATA (Roe, 1998), 'the' is given a frequency score of 609.35 per 10,000 words – or 6%.

as inherently 'given'." (Martin, 1992, p.122). That is, they do not need to be invoked even initially for the writer to assume that the entity is already in 'the air', or the mind, of the reader. For example, if there is only one of something in the context of culture, such as the sun, the moon or the Atlantic ocean, the writer can assume familiarity. When the 'location' of the identity of the element is to be found in the 'context of culture', it is known technically as 'homophoric' reference. For Martin (1992), the context of culture

> embraces relevant information which cannot be perceived, but which can be assumed because of shared knowledge among interlocutors deriving from their membership in some definite community. (p.121)

While psychological and computational theories have failed to identify clearly the extent of presuppositions that may pertain at any discourse moment (see sections 2.2.2.5 and 2.2.4), this description of 'some definite community' is also open to a variety of interpretations and needs to be refined to be considered reliable. Martin's scheme allows the speaker to appeal to a shared social context as the source of distinction between participants, but dependent on the speaker's context of culture, which is relative and can vary according to the speaker's intentions.

Fig. 2.8 shows the analytical choices available for a referent up to the selection of homophoric reference. If the identity of the referent is not retrieved from the context of culture, it is retrieved from the context of situation. When the interpretation of an entity is related to an item in another sentence, a cohesive tie is created. A cohesive tie contributes substantially to the texture of a text "since it creates a network of lines of reference, each occurring up to and including the initial reference." (Martin, 1992, p.52) Clearly, then, there must be a point at which the participant is mentioned for the first time. If the identity of a participant is not expected to be retrieved from the context or co-text – ordinarily if the participant is presented – then in the Tracking network this can be termed 'Addition' (Martin, 1992). Although there has been no previous mention of this participant, and so participants with Addition reference do not create co-referential cohesive ties with any previous participant in the text, an Addition-participant provides the potential to be tracked later in the text – it is an 'anticipant' waiting to become focalised and pronominalised (see section 2.7.6).

Martin's (1992) analysis of phoricity is, in his own terms, synoptic (see also section 1.2.2). Although Martin distinguishes between a phoric and cohesive analysis of text in terms of dynamism, thus:

> these retrieval categories broach upon a dynamic as opposed to synoptic perspective on participant identification ... The choices reviewed are relevant both to decisions a speaker must make when selecting phoric or non-phoric items and to processes the listener must go through to recover any information that is presumed. (p.126)

the phoric system is not represented dynamically, resulting in a synoptic approach to a dynamic feature of text. This recognition of the processes employed by speakers/writers and listeners/readers echoes the concern expressed by Emmott (1994; 1997, see section 2.7.5) regarding the more dynamic and prospective nature of anaphora.

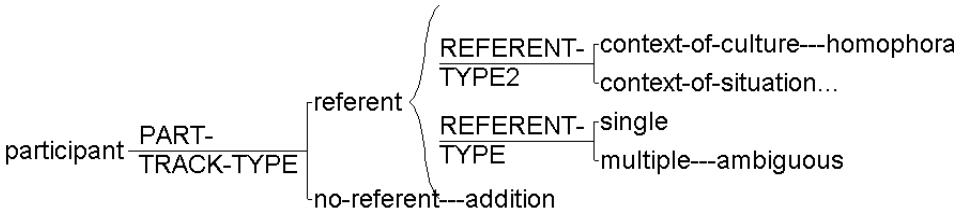


Fig. 2.8 Phoricity Network up to Homophora

Both Identification and Tracking depend on the same unit – that of the Participant – and so they inevitably converge. Addition is associated with the choice of Presenting reference and Presuming reference is associated with choices beginning at the point of 'Referent'. The entry condition for the phoricity system appears to be 'Specification' in the Identification system; a participant cannot be tracked if it has not been identified. Generic implies Presuming homophoric reference – knowledge located in the context of culture (Martin, 1992).

Fig 2.10 shows the possibility for the interpretation of a referent to be ambiguous in a text. Ambiguity arises when interpretation of an entity cannot be tied to a single entity. Ambiguity can occur regardless of the location of the identity of a referent; the ambiguity system is parallel to the choice between the contexts of situation and culture. Where there are multiple referents, ambiguity potentially disrupts the flow of communication.

2.4.2.2 Exophoric and Endophoric Reference

If the entity is not retrievable from the context of culture, it may be retrieved from the context of situation, defined as "relevant information that can be perceived… including text" (Martin, 1992 p.121). When the identity of a participant is retrievable from the context of situation, the reference is either Exophoric (retrievable from a non-verbal source) or Endophoric (retrievable from within the co-text). This distinction is illustrated in Fig. 2.9. An Exophoric reference to an entity is 'outside' the co-text. As an exophoric

participant is tracked through a text, it may be re-invoked from the non-verbal context, or subsequently considered part of the endophoric co-text, or both. In this study the option that analyses the link as exophoric is used.

If the identity of the Endophoric referent is related to a point earlier in the co-text, the relationship is anaphoric, or "backward-pointing". If the identity of the participant is to be retrieved from an entity after mentioning but not identifying it, for example when a pronoun precedes its referent, the reference is following ("forward pointing"). Martin (1992) distinguishes forward-pointing reference within the same (nominal) group, which he labels esphora, from the less common forward-pointing reference across groups, for which he reserves the term cataphora. This completes the grammatically-based resources used for indicating phoricity. Concurrent with the final grammatical resource for retrieval of participant identity is the beginning of the more lexically-oriented resources for referring to the same participant. These resources will be reviewed in the next section, and discussed further in section 2.4.2.4.4.

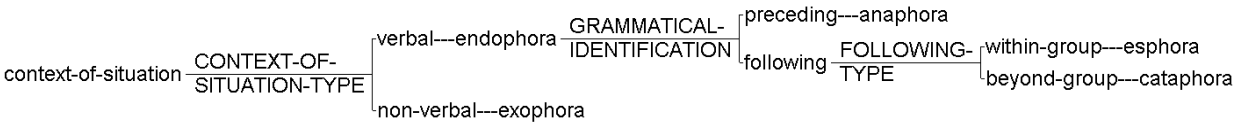


Fig. 2.9 Phoric Relations within the Context of Situation

2.4.2.3 Direct and Indirect Reference

Within endophora, there is both a grammatically-signalled system, discussed above, and a more lexically-based system for Participant Tracking. That is, the identity of a participant is retrieved by using lexical relations in parallel with grammatical relations. In Martin's retrieval network for participant Identification and Tracking, the final distinction between direct and indirect reference is simultaneous with the 'direction' of phoricity within the text, as illustrated in Fig. 2.10.

Martin (1992) divides the lexical resources into Direct and Indirect. Direct reference retrieves participant identity and creates cohesive ties through complete or scattered repetition, the latter being divided into inflexion (same word class, *e.g. win-won*) and derivation (change of word class, *e.g. win-winner*). Indirect reference covers the implied relationships labelled *bridging* by Haviland and Clark (1974; Clark and Haviland, 1977; Clark 1977) to mean items that "may presume information that is implied rather than

directly retrievable" (Martin, 1992, p.124). (See section 2.2.2 for discussion of 'Bridging'.) Martin lists typical implicit lexical relations between the participant and its dependent referent, including part-whole, class-subclass, and experiential relationships (which are more collocational in manner), and notes that "Bridging depends on experiential connections between presuming and presumed which facilitate the recovery of an implied identity." (1992, p.124).

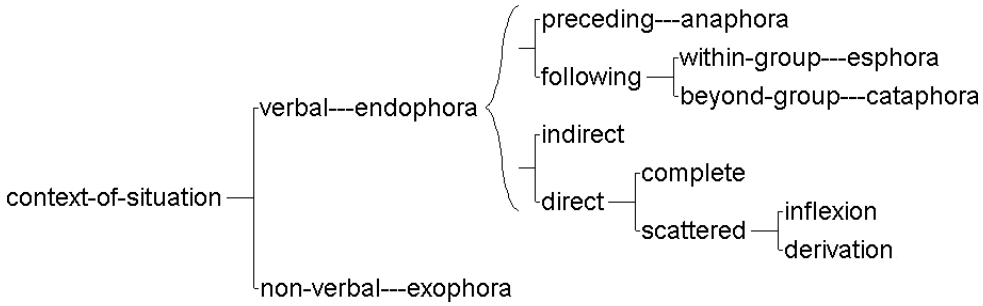


Fig. 2.10 Endophoric Reference

These experiential connections will now be examined, with the intention of integrating them within the analytical model. As "bridging may depend on experiential relations of various kinds" (Martin, 1992, p.125), it is important to identify these relations. Martin (1992) explores the Experiential metafunction and distinguishes lexical repetition with derivation of a single lemma from other less explicit lexical relations. Martin also notes the developments made in Halliday and Hasan (1985) (reproduced as Table 2.2). For the purposes of participant tracking, all of these relations of co-reference, co-classification and co-extension are important as they all require the participant to be a part of the local co-text. Co-classification and co-extension re-invoke the presumed participant, while tying it with another presented or presumed participant to create a cohesive relation of Participant Tracking. Thus, the participant is necessarily a part of the reference, even though it may not be co-referential. In Martin's (1992) terms, they are still Reminding reference.

More importantly for this study, it is typical for lexical and grammatical relations to support each other in text. Thus, while it is possible to identify lexical cohesive ties between participants using lexical relations, the only lexical ties that will be analysed in this study will be those that are grammatically signalled for Participant Tracking. That is, items must have grammatical cohesive ties, at which point lexical ties will also be examined; lexical reference and grammatical reference are hypothesised to be simultaneous systems. Grammatical reference is required in the analysis because there are almost inexhaustible

possibilities when looking for lexical relations within a text, many of which cannot be reliably tied to other items. Consequently, only those lexical relations that are signalled as being phoric will be analysed. In this way, the analysis will track participants through relations of co-reference (Table 2.2), rather than attempt to trace participants through a larger lexical cohesion analysis (Hoey, 1991b). For this purpose a detailed view of lexical relations can be employed. These relations are the focus of the following section.

<table>
<tr><td colspan="3">NON-STRUCTURAL COHESION
COMPONENTIAL RELATIONS</td><td>ORGANIC RELATIONS</td></tr>
<tr><td rowspan="3">Grammatical Cohesive Devices</td><td>Device</td><td>Typical Tie Relation</td><td rowspan="3">Conjunctives
e.g. causal tie
concession tie ..

Adjacency Pairs
e.g. question
(followed by) answer;
offer (followed by)
acceptance);
order (followed by)
compliance</td></tr>
<tr><td>A. Reference
1. Pronominals
2. Demonstratives
3. Definite article
1. Comparatives</td><td>Co-reference

Co-reference &
Co-classification</td></tr>
<tr><td>B. Substitution & Ellipsis
1. Nominal
2. Verbal
3. Clausal</td><td>Co-classification</td></tr>
<tr><td rowspan="2">Lexical cohesive devices</td><td>A. General
1. Repetition
2. Synonymy
3. Antonymy
4. Meronymy

B. Instantial
1. Equivalence
2. Naming
3. Semblance</td><td>Co-classification
or
Co-extension

Co-reference
or
Co-classification</td><td>Continuitives
(e.g. still, already …)</td></tr>
<tr><td colspan="3">STRUCTURAL COHESION

A. Parallelism
B. Theme-Rheme Development
C. Given-New Organisation</td></tr>
</table>

Table 2.2 Summary of Cohesive Devices (Halliday and Hasan: 1985 p.82)

2.4.2.4 Lexical Relations

As lexical items accumulate in a text to a 'critical mass' to reveal the Contextual Configuration and register (Halliday and Hasan 1985) of a text, it is extremely likely that lexical items create a host of cohesive ties, both explicit and implicit, contributing to the texture and consistency of a text. The context of situation necessarily restricts the lexical choices available. Entities in a text are tied in lexical cohesive relations, but not all those relations are co-referential or are related to (potential) participants, essential in participant tracking. Because this analysis focuses on participant Identification and Tracking in order

to examine the contribution that the separate textual systems of Reference, Theme and Information to the clause, it is unnecessary to analyse all cohesive relations.

Hoey, for example, provides a detailed schematic (1991b, p.58-60), to determine whether lexically related items are co-referential, and discusses how lexical relations work to co-refer. This analysis is likely to exhaust all cohesive lexical relations within a text, and more than compensates for Halliday and Hasan's (1976) emphasis on the grammatical aspects of cohesion despite their assertion that "Some forms of cohesion are realized through the grammar and others through vocabulary" (Halliday and Hasan, 1976, p.6). As Hoey (1991b) demonstrates, lexical cohesion is more prevalent in text than grammatical cohesion. In this study, however, cohesion analysis is being employed only so far as it is useful to track participants. Consequently, using the network discussed so far, the grammatical resources for co-reference will be used as the starting point for participant tracking because the grammatical relations force the reader to see the lexical ties as co-referential. That is to say, the grammatical resources of Participant Tracking make lexical ties explicit so that the reader's intuitions about lexical cohesion are confirmed. Without grammatical signals, each reader's map of lexical ties is likely to exhibit considerable variation.

While cohesive ties created by lexical relations contribute to coherence, while contributing to the Contextual Configuration, it is only when grammatical relations force the reader to consider that the items are tied that participant Tracking can take place (Halliday and Hasan, 1985 p.82). The most obvious example of this is when a lexical item is repeated with Presuming reference compared with lexical repetition with Presenting reference. Following on from Halliday and Hasan's (1985) categories of lexical relations, Martin (1992) develops a more comprehensive system using Lyons's (1977) categories of semantic relations. The full classification of lexical relations is illustrated in Fig. 2.11, with typical realisations, and uses Martin's terms wherever they differ from Lyons.

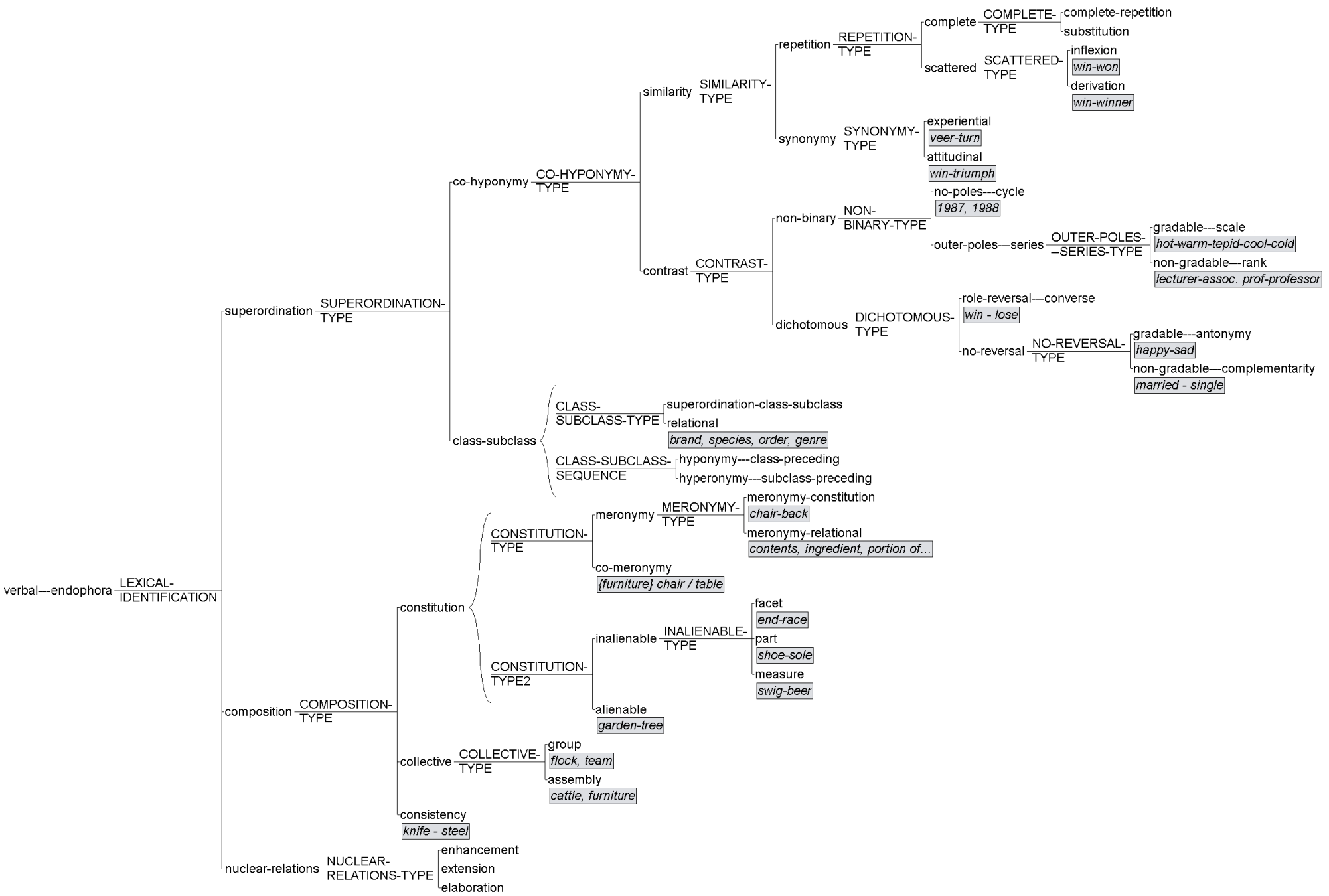


Fig. 2.11 Network of Lexical Relations for Participant Tracking (from Martin, 1992)

The main differences between the model presented here and Martin's (1992) version derive from applying the model to participant tracking, and so simplifying the model to reduce the number of simultaneous systems. In Fig. 2.10 (above), grammatical and lexical systems are shown in parallel, but the Direct and Indirect lexical options for tracking are shown as alternatives. As the lexical relations described here are intended to define or replace the Direct/Indirect distinction, and only one lexical relationship is required to establish a cohesive connection, the lexical choices are considered as alternatives. For instance, Superordination and Composition are presented in Martin (1992, p.294) as simultaneous systems within taxonomic relations. Because this study only requires one relationship in order to test the validity of the network to track participants through simultaneous grammatical and lexical networks, the analytical network does not require both options, although more than one relationship may indeed be possible. Similarly, Nuclear Relations are included here as an alternative to Superordination and Composition rather than as a simultaneous system. Martin (1992, p.294) includes Activity Sequence as part of the network of semantic choices, but since these relations refer largely to inter-clause aspects of cohesion, the only Activity Sequence system that has been included is Nuclear Relations because these can be applied Participants.

The main classifications of lexical relations – Superordination, Composition, and Nuclear Relations (see Fig. 2.13) – are detailed below. Compared to Halliday and Hasan's (1985) revision of the taxonomy of cohesive lexical relations in Halliday and Hasan (1976), Martin's Superordination details Halliday and Hasan's (1985) Repetition, Synonymy (including hyponymy), Superordinate and General Item, while Halliday and Hasan's (1985) Meronymy is just one of the relations subsumed under 'Composition' relations. Martin (1992) gives Collocation more detail than in Halliday and Hasan (1976), including it in 'Nuclear Relations' (see section 2.4.2.4.3), which are based on logical relations between components within the clause.

2.4.2.4.1 Superordination

Within Superordination the main distinction is between Co-hyponymy and Class-subclass; that is, whether the relation between two lexically-cohesive items is of the same order or of classification, respectively (Martin, 1992, p.301). Co-hyponymy is divided into Similarity and Contrast. Within Similarity, Repetition can be Complete or Scattered either through Inflexion (using the same word class) or Derivation (changing word class). Substitution, the alternative to Complete Repetition, is most commonly realised by pronominalisation,

and has been added to the system network for this study to complete the options for Repetition. If the Similarity is not Repetition, the other choice is between Experiential synonymy, (another item with similar Experiential content, *e.g. veer* for *turn*) and Attitudinal synonymy (another item which changes the Interpersonal content while retaining the experiential content, e.g. *triumph* for *win*). Although there are other relations subsumed under the term Superordination, and Martin's (1992, p.294*ff*) network looks a little different to Fig. 2.13, it must be remembered that the network being proposed here extends only to lexical relations of cohesion, not to all experiential lexical-semantic relations.

The term 'Similonymy' was proposed by Bawcom (personal communication) to show that there will never be exact similarity between alternative terms, because their collocational patterns will vary (Sinclair, 1991), preventing complete synonymy, and because referents pass through various transitive processes, thus transforming their experiential content (Brown and Yule, 1983, p.201) and precluding true repetition. Thus, Similonymy would be a more appropriate term to apply at the 'Similarity' point in the network of lexical relations. The lack of complete synonymy in general lexical relations also explains why Halliday and Hasan do not consider that alternative lexical items can realise co-reference. However, I would suggest that all relations of Similonymy can be used in a text to instantiate relations of co-reference.

On the same scale in a taxonomy, and so considered co-hyponymous, Contrast lexical relations divide into non-binary and dichotomous relations (*cf.* Jones (2002) who prefers the term antonymy for all senses of Contrast or 'oppositeness'). Non-binary relations divide into Cycle (without limit, typically a numerical series), and Series, which subdivides into Scale (a gradable series where items run into each other) and Rank (discrete items) (Martin, 1992, p.302*ff*). Dichotomous relations divide into Converse, showing role reversal (where one item necessarily entails another, *e.g. winner-loser*), and No reversal which subdivides into Antonymy (which is gradable, as Scale, *e.g. happy-sad*) and Complementarity (which is ungradable, and also exclusive such that only one can pertain at one time, e.g. *married-single*). Whether they contrast or are similar, all co-hyponymous relations can provide alternatives for a lexical item on the same scale within a taxonomy. The next distinction differs in that the potentially co-referential lexical items are of a different order.

Within Class-Subclass, there are two simultaneous systems (Martin, 1992, p.299). The first is related to the type of relationship between the two terms. If one of the items names a general term, such as *brand* or *species*, it is Class-Sub-class – Relational. Otherwise, the item is termed Class-Sub-class. The other simultaneous system identifies the sequence in text of the tied participants: Hyponymy is when the class precedes the member (*e.g. flower-rose*), and Hyperonymy is when the member precedes the class (*e.g. rose-flower*).

2.4.2.4.2 Composition

Relations of Composition are divided into Constitution, Collective and Consistency, the latter representing a 'made of (material)' relationship that has no further levels of analysis (e.g. the *leather* of a *shoe*). Collective is divided into Group and Assembly, "both of which have the function of treating collections of individuals as aggregates" (Martin, 1992 p.306), the difference being that Group operates as a pre-numerative ('a flock of'), while Assembly does not.

Constitution is divided into the two simultaneous systems of Meronymy and Alienability. Items with relations of Constitution can be Co-meronymous (both forming part of a larger whole, *e.g. monitor, keyboard*), or may be related by Meronymy (one is a part of another, *e.g. cpu, motherboard*). If one of the ties names the relationship, e.g. *ingredient*, *portion*, *part*, then the relationship is labelled Meronymy-Relational, if not the relationship is labelled Meronymy (Martin, 1992, p.304). Within Constitution, Choices in Alienability are concurrent with Meronymy, and relate to the separability of the ties. Alienable items are exemplified by a pair such as *tree* and *garden*. A tree need not be a part of a garden, and a garden need not include a tree. Inalienable pairs are necessarily tied in relationships of Facet (*races* have *a finishing line*), Measure (a *swig* of *beer*) or Part (*sole* of a *shoe*) (Martin, 1992, p.305).

2.4.2.4.3 Nuclear Relations

The final type of lexical relationship is an attempt by Martin (1992) to locate collocational relationships within a metafunctional and discoursal framework. Halliday and Hasan (1976) contrast collocational cohesion with forms of reiteration, and classify it as "a cover term for the cohesion that results from the co-occurrence of lexical items that are in some way typically associated with one other, because they tend to occur in similar environments" (p.287). Using the metafunction of logical relationships, Martin identifies

the nuclear relations of Extension, Enhancement and Elaboration as essential to collocation.

Nuclear Relations configure participants and processes as Activity Sequences; they "reflect the way in which actions, people, places, things and qualities configure in activity sequences" (Martin, 1992p.309). Because people, places and things are often realised as Participants, Nuclear Relations may contribute to tracking Participants in discourse. Nuclear Relations describe typical configurations of Participants and Processes that may develop in a text as a result of activity sequences.

The main nuclear relations are exemplified in Table 2.3, which details how the different logical relations are realised by clausal, nominal and verbal roles. Collocational ties create cohesion and texture, but are rarely used for co-reference. However, through lexical relations it is possible to co-refer using nuclear relations. That is, by using a grammatical marker to suggest a co-referential link between two lexical items linked, for example, through the same process, the reader may use a nuclear relation to track the two participants.

	Elaboration =	**Extension** +	**Enhancement** x
Clause	PROCESS = RANGE: PROCESS take shot (*take a shot*)	PROCESS + MEDIUM + RANGE: ENTITY shoot deer (*shoot the deer*)	PROCESS X CIRCUMSTANCE shoot field (*shoot in the field*)
Nominal group	CLASSIFIER = THING practice shot (*a practice shot*)	EPITHET + THING loud shot (*a loud shot*)	THING X QUALIFIER shot dark (*a shot in the dark*)
Verbal group	EVENT = PARTICLE shoot up (*shoot up*)	EVENT + EVENT try shoot (*try to shoot*)	EVENT X QUALITY shoot carefully (*shoot carefully*)

Table 2.3 Elaboration, Enhancement and Extension across clauses and groups (adapted from Martin, 1992, p.317)

The example below provides two examples of nuclear relations

```
The provision of mechanisms for the protection of data and
other computer-based resources and for securing networked
transactions is the concern of this chapter.
```

The provision can be analysed as having a Nuclear Relation with the endophoric *mechanisms* through the relation of Extension; the relation between *mechanisms* and *provided* is typically that of Process and Range, but these have been changed through a process of grammatical metaphor (where the typical grammatical role has been changed to a different role in the clause, such as when a verbal process becomes nominalised). The same analysis can be applied to *securing* and *networked transactions* (itself another grammatical metaphor).

2.4.2.4.4 Discussion

The categories above (section 2.4.2.4.1 to 2.4.2.4.3) have been included in order to make explicit the semantic relations that operate between lexical items with phoric reference. They encompass the system of relations that enables direct, indirect or bridging reference, once the text has established a grammatical relationship. Returning to the first label within these lexical categories (Repetition), it should be clear that the choices here are identical to those at the entry point for Direct reference (see Fig. 2.12 and 2.13). That is, Direct reference relies on lexical repetition and pronominalisation, while Indirect reference, or Bridging, depends on the range of semantic, experiential relations detailed above. Since the experiential network places lexical repetition and pronominalisation on the same cline as other lexical relations (Martin, 1992, p.301), there seems little reason to maintain the distinction of Direct and Indirect reference, particularly considering the theoretical status of Bridging, as discussed in section 2.2.2. This fusing of the grammatical and lexical resources of the language is to be expected considering the assertion in SFL that lexis and grammar are distinguished by degree, or delicacy, not type. The scheme being hypothesised here, in fact, sees the 'crossover' point between lexis and grammar as the point where morphology meets semantic relations. In the lexical options we see that Derivation is hypothesised to be closely related to Synonymy.

The choices in Fig. 2.13 are not randomly arranged. As in all of the system networks in this study, the choice that is hypothesised to be more frequent, Unmarked, or the default, is placed at the top for each choice, as suggested by Fawcett (1988 p.204). Presuming, for instance, is placed above presenting because it is hypothesised to be more frequent in text (see Appendix 2.1). That is, the order of lexical relations is hypothesised to run from unmarked at the top to marked at the bottom. For instance, it may be more common to connect an adjectival derivative to a nominal lexical item than to its nominal antonym because relations of Similarity are easier to access than relations of Contrast. Hence, Derivation is above Antonymy in the network. As we move down the hypothesised scale,

the connection between items becomes more implicit, and *bridging* may take place, if bridging means searching for semantic relations. The relations are ordered with the hypotheses that, first, each terminal point in the network will be more frequent than the one below, and, second, that each terminal point in the network will take slightly more time to for a language user to 'process' than the one above: "the vertical order of features in a system … could in fact be put to work to carry the meaning 'Try the features in this sequence'." (Fawcett, 1988 p.204). These hypotheses clearly need to be confirmed by comparison with corpus results or to psycholinguistic tests. An attempt is made to test the first hypothesis (see section 2.6), at least within the registers examined here. The second hypothesis is beyond the scope of this study, but can be implied from the first according to a usage-based model of language (Bybee, 2006; Ellis, 2006; MacWhinney, 1998). If the hypothesis is confirmed, it may account for some anomalous data in the Bridging experiments (see discussion in 2.2.2).

Lexical items frequently co-refer in more than one way. That is, a participant may co-refer with another entity through both co-meronymy and co-hyponymy: contrast, or it may co-refer with more than one other participant (*i.e.* it is *ambiguous*) through more than one semantic relation. The network is ordered so that the relationship which is hypothesised to be the most congruent (complete lexical repetition) is at the top, with the least congruent (nuclear relations: elaboration) at the bottom. An alternative solution, rejected in this study, would be to provide the 'null-option' (no Superordination, no Composition, and no Nuclear Relations) and allow the three categories of lexical relations to become simultaneous systems (as in Martin's original model). This merits further investigation. However, since the aim here is to be as conservative as possible in proposing lexical relations, only the relation hypothesised to be the easiest to process, because it is the most common, will be analysed. For the time being, it is only sufficient to demonstrate the applicability of the model. Refinements, such as simultaneous systems, can be investigated in further studies.

One of the clearest differences between Martin's (1992) scheme for participant tracking and Halliday and Hasan's (1985) outline for lexical cohesion (see table 2.1 above) is that Martin does not consider Instantial lexical relations to be a distinct option within the network. Instantial and General lexical relations can be seen as the poles on a cline, rather than as distinct options. General lexical items represent meanings that are culturally commonly associated and learned through frequent exposure and association. In some texts lexical relations will be made explicit to the reader, and can be seen as instantial,

while those same items and relations will remain implicit in another text and therefore be considered as General. For example, a technical text written by a specialist for peers will assume a context of culture that contains a wide range of technical lexis, which will appear in the text with General lexical relations. The same writer on the same subject writing for a lay audience will then need to make those same general lexical relations more explicit by changing them into instantial relations. Thus, General lexical relations are context-dependent, and may not be as 'General' as they appear.

A 'General' system of lexical relations may also suggest a fixed context-free view of meaning. This position will not be supported in this thesis, as an attempt will be made to incorporate a constructivist view of cognition into discourse processing and to demonstrate how the textual metafunction contributes to the construal of textual experience through meaning (Halliday and Matthiessen, 1999). As an example of this perspective, Halliday and Matthiessen (1999) are quite clear when they say

> it is the grammar that construes the experience, that constructs for us our world of events and objects … Meanings do not 'exist' before the wordings that realize them. They are formed out of the impact between our consciousness and its environment. (p.17)

A model of reference that depends on the meanings of referents to inhere in an external 'database' or memory cannot account for novel meaning-making. Alternatively, a constructivist model enables meanings to be construed instantially, with conventional meanings acting as examples of socially-accepted categories rather than representing an idea that exists in the world or in a memory (Thibault, 1999; 2004a). As a constructivist approach to meaning in text will be an important aspect of the current study (section 1.2.4), studies which depend on a representationist model of meaning must be critically evaluated before they can be incorporated into any proposed model.

A register- and context-dependent approach to language supports the notion of a discourse community (Swales 1990) and a probability approach to Register (Halliday, 1991; Halliday and James, 1993; Moore, 2006; Nesbitt and Plum, 1988). For this study, there will be no distinction between instantial and general relations, particularly as the majority of the texts will be designed to induct readers into a discourse community, and so the distinction between the two will not only be difficult to prove at times, but may prove inconsequential, especially for the analysis of participant identification.

2.5 Analytical Model

The remaining sections in this chapter use the preceding theoretical review and discussion to analyse texts (itemised in Table 1.2). This section summarises the model, and outlines the methodology of the analysis.

The model used in this analysis is based largely on Martin (1992). Having discounted a range of studies that appear to be studying similar linguistic features, it is clear that only a limited amount of research can be drawn directly into the same framework, due to methodological approaches (see sections 2.2.2 and 2.2.3) or differences in the model of language and its role in cognitive processes (see section 2.4.2.4.4). Very little has been changed in Martin's (1992) scheme, except for the greater delicacy in analysis offered by detailing the types of relationships labelled 'bridging', but here expanded into the full range of experiential relations.

One reason that the taxonomy of experiential relations combines so easily with the system network for participant Identification is the way that Martin (1992) uses system networks as taxonomies for the analysis of discourse semantics. It is possible to approach system networks as representing either a constitutive or a compositional taxonomy. Thus, in some system networks for Theme, the entry condition is a clause, which *consists of* a Theme and a Rheme, and the Theme *consists of* Marked or Unmarked Theme, and so on. In Martin's (1992) networks of participant Identification and Tracking, the relationships between choices in the networks are not constitutive, but classifying. That is, typical of taxonomies, once a unit has been decided upon, the network allows the analyst to classify the function of that unit. For example, once a participant has been identified, it is *classified as* having the function of Effected or Neutralised, if Effected, it is *classified as* having the function of Specified or Generalized, and so on (see Fig. 2.3). This approach will be followed in all the system networks used for analysis in this study.

The model of analysis that will be used in this section of the study has two parts. The first – participant Identification (appendix 2.1) – distinguishes participants and enables their tracking through text. The second model of analysis is that of participant Tracking, (appendix 2.2), which depends on phoric relations between participants. These two systems have obvious points in common. Most importantly, they have the same unit of analysis – the participant in ergativity. Consequently, texts were segmented first by group

(see section 1.2.3), and then only those groups that are participants or potential participants (see section 2.1 and 2.4) were analysed. A decision was made to 'parse' groups in the analysis depending on whether they could be identified as being modified or not, following the argument for 'of' as a "postposition" in Sinclair (1991). That is, a dynamic, syntagmatic analysis was adopted such that a group was defined up to the point where it became clear whether the group had been completed. Consequently, verbal groups retained their adverbial particle, including adverbial particles used to introduce an indirect object, and nominal groups retained their post-modifying preposition. Thus, the example sentence below (Table 2.4) could have stopped at 'provided'. The reader knows that more is to come when 'through' is given. Other examples include 'and' after 'leadscrew' and 'of' after 'front'. For participant analysis this has the advantage that most (potential) participants are isolated and commence with their deictic realisation for Identification and Tracking, although it creates groups that do not match more traditional categories such as prepositional phrase or adverbial group.

Alternative hand movement	**is provided through**	**a leadscrew and**	**nut**	**by a handwheel at**	**the front of**	**the knee.**
No indication of further modification	Verbal group with modification	Nominal group to be post-modified	No indication of further modification	Modification separated from its verbal group, with further modification	Nominal group to be post-modified	No indication of further modification

Table 2.4 Example of division between groups depending on possibility of post-modification

Dividing nominal groups in this way also has the advantage that both Head and Thing, where they are different, can be analysed for reference. In many nominal groups the experiential Thing and the logical Head are conflated. In the group *the integrity and privacy of information and other resources* (Sec) there are two Heads, *integrity* and *privacy*, which are post-modified. Logically, they are Head because all other parts of the group are dependent on them, and because "It is the Head that determines the value of the entity in the mood system, and therefore as potential subject." (Halliday and Matthiessen, 2004 p.333). The participants *integrity* and *privacy* are also Thing because they are the main participant "functioning in transitivity" (p.333) and are the "semantic core of the nominal group" (p.325). However, in come cases the same structure may be interpreted such that the Head and Thing are separate. The Thing of the group may be premodified by the Head as a deictic (*e.g.* "another one of my friends"), a numerative (*e.g.* "three of those tiles"), an

epithet (*e.g.* "that monster of a skyscraper") or a classifier "where the Head word specifies the class to which the Thing is said to belong" (Halliday and Matthiessen, 2004 p.333). For instance, in *a factual statement of physical phenomena* (EAP) the Thing is *physical phenomena* which is premodified by the Head *a factual statement* operating as a classifier – it tells you that in this case *physical phenomena* belong to the class of *a factual statement*. Verb-subject agreement remains with *a factual statement* but we are discussing one class of *physical phenomena*.

The two analytical networks of participant Identification and Tracking were applied to sample texts (section 1.3.2). Typically, certain choices coincide (e.g. Presuming coincides with anaphoric reference), but rhetorical effect can be achieved by not aligning the two systems. That is, grammatical items indicate participant tracking, but no previous mention, or likely lexical relation is apparent in the co-text, or Presenting reference can be used with a participant that has already been identified. The texts were analysed independently for features in the two systems, and the results were drawn together at the end of the analysis. This procedure was followed so that results from the two systems would not interfere with each other. The tools used in the analysis are described below. The results, and difficulties encountered in analysis, will then be discussed in section 2.6.

Although the analytical models do not specify a view of cognition, they must be compatible with textual and cognitive processes that allow text to construe and accrue meanings without the constant need to physically or mentally search back in text for co-referents (see section 2.7). The model and tools used in this analysis do not contradict such a view. Although the model does not track participants dynamically, it does not preclude such a model and its implications for cognitive, and computational, processing. This aspect of this analysis requires development in the future if it is to become more responsive to a dynamic approach.

The model of participant Identification and Tracking identified through the previous discussion, and illustrated in appendices 2.1 and 2.2, was integrated into the main tool used in this study – the software packages of Systemic Coder (O'Donnell, 2003) and UAM Corpus Tools (O'Donnell, 2007). The software allows the analyst to choose their own network and labels. In this case, the networks described in section 2.4 were 'programmed' into the software. (Most of the illustrations of the systemic networks within this thesis were also developed using this software). The analyst then uses these networks to manually code, or tag, texts which have been segmented into units of analysis

determined by the analyst (here, participants). After the text has been tagged, the analyst can then test the results, using the descriptive and comparative statistics included in the software, on their own or in combination with other data sets using the same network. Alternatively, the results can be exported to other statistical packages, such as SPSS or Microsoft Excel. Perhaps the greatest drawback of Systemic Coder is that it does not allow multiple analyses of the same unit. Consequently, recursive, multiple or ambiguous choices could not be coded simultaneously. Only one analysis was selected. Although the later releases of UAM Corpus Tools did allow for multiple coding for a single segment (as demonstrated in Chapter 3), it was decided that a single-layer sequential, rather than hierarchical, analysis was most suitable for a dynamic approach.

As the results in this part of the study are mainly descriptive in nature, Systemic Coder (v2.0) proved sufficient for most of the participant analyses. Appendix 5 provides a copy of the UAM programme for installation and all of the results for this project so that the reader can examine all choices made in the textual analysis.

2.6 Participant Identification and Participant Tracking Analyses

The analyses are presented below as an illustration of the applicability of the model. As well as presenting a quantitative analysis, mainly descriptive in nature, the discussion will focus on textual issues, and observations regarding typicality of features for this genre. Issues and difficulties encountered in the analysis of texts are discussed in sections 2.6.1.2 and 2.6.2.2. The quantitative results are meant only as an initial test of the models. The results obtained are not intended to be directly extrapolated to the language in general, although the chosen texts are considered to be typical of this register. Examples of all categories in the Participant Identification analysis of text AN are illustrated in appendix 2.3, and examples of all categories in the Participant Tracking analysis of text MM are illustrated in appendix 2.4. Appendix 5 contains all analyses and examples.

2.6.1 Participant Identification

The system of Participant Identification enables the reader to distinguish between different participants in the text. This is achieved mainly through reference, especially through choices in Presenting and Presuming, which tells the reader that the participant is new to the discourse or that the identity is to be located in the context. The full

Identification system consists of the three simultaneous choices of Presenting-Presuming, Comparison, and Specific-Generic. There is no reason to assume that the first two systems are likely to exhibit atypical results, but the frequency of Generic is expected to be greater than in non-instructional text (Martin, 1989).

2.6.1.1 Typical Analyses and Realisations

Fig. 2.12 illustrates the likelihood of each choice in the network in the texts analysed (see section 2.4.2.4.4). The number and percentage of features in the identification system are detailed in Table 2.5. The following is a commentary on these results.

All but 62 (2.6%) of the 2321 participant groups are Effected in the Identification system. In all texts the Neutralised choices are in titles. In BN (see table 1.2 for all text acronyms), there are also Neutralised choices in bulleted lists of items. The system of Participant Identification is Effected in the vast majority of cases. Identification, therefore, plays a significant part in developing the textual meanings of these texts. That is, the text would be unable to develop meanings very effectively without this system.

Of 2321 effected groups, only 16 are Generalized (section 2.4.1.1). Generalized is realised by *it*, *e.g.* "it pays to begin with a system that is free of preventable interference" (AN), *there*, ("unless there is a policy for its use" from BN), *nobody*, and various fixed phrases including *for example*, *in service* and *by chance*. Although they resemble nominal groups, none of these groups can be identified or tracked. The remaining 2305 groups that enter the main choices in the Identification system are fairly evenly split between Presenting (46%) and Presuming (54%).

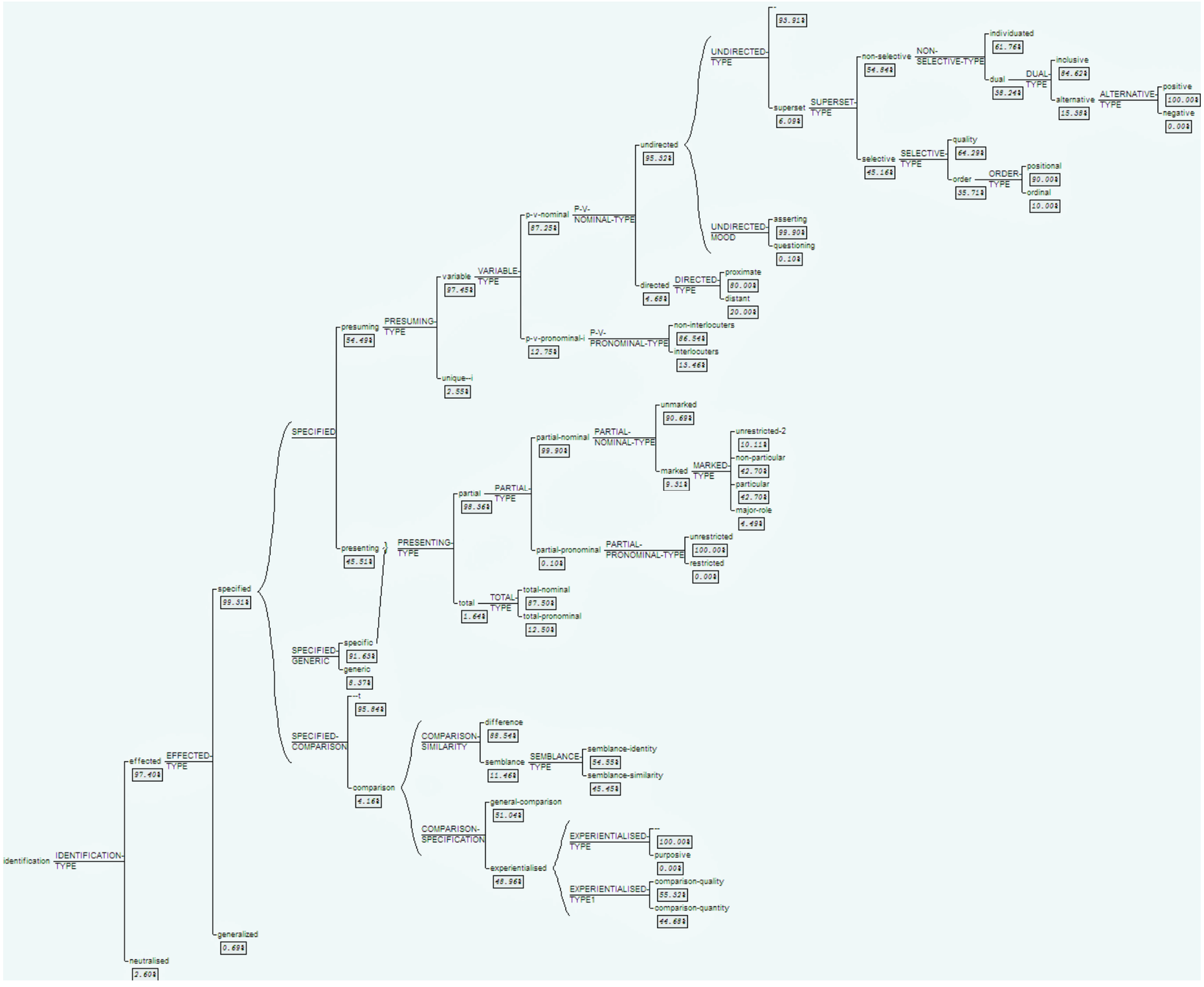


Fig. 2.12 **Selection of Options, as Percentage, of Participant Identification for Combined Texts**

Feature		Mean	N
IDENTIFICATION	effected	97.40%	2321
	neutralised	2.60%	62
EFFECTED	specified	99.31%	2305
	generalized	0.69%	16
SPECIFIED	presuming	54.49%	1256
	presenting	45.51%	1049
PRESUMING	variable	97.45%	1224
	unique--i	2.55%	32
VARIABLE	p-v-nominal	87.25%	1068
	p-v-pronominal-i	12.75%	156
P-V-NOMINAL	undirected	95.32%	1018
	directed	4.68%	50
UNDIRECTED	-	93.91%	956
	superset	6.09%	62
SUPERSET	selective	45.16%	28
	non-selective	54.84%	34
SELECTIVE	quality	64.29%	18
	order	35.71%	10
ORDER	ordinal	10.00%	1
	positional	90.00%	9
NON-SELECTIVE	individuated	61.76%	21
	dual	38.24%	13
DUAL	alternative	15.38%	2
	inclusive	84.62%	11
ALTERNATIVE	positive	100.00%	2
	negative	0.00%	0
UNDIRECTED-MOOD	asserting	99.90%	1017
	questioning	0.10%	1
DIRECTED	proximate	80.00%	40
	distant	20.00%	10
P-V-PRONOMINAL	non-interlocuters	86.54%	135
	interlocuters	13.46%	21
PRESENTING	partial	98.36%	957
	total	1.64%	16
PARTIAL	partial-nominal	99.90%	956
	partial-pronominal	0.10%	1
PARTIAL-NOMINAL	unmarked	90.69%	867
	marked	9.31%	89
MARKED	unrestricted-2	10.11%	9
	non-particular	42.70%	38
	particular	42.70%	38
	major-role	4.49%	4
PARTIAL-PRONOMINAL	unrestricted	100.00%	1
	restricted	0.00%	0
TOTAL	total-nominal	87.50%	14
	total-pronominal	12.50%	2
SPECIFIED-GENERIC	specific	91.63%	2112
	generic	8.37%	193
SPECIFIED-COMPARISON	--t	95.84%	2209
	comparison	4.16%	96
COMPARISON-SIMILARITY	difference	88.54%	85
	semblance	11.46%	11
SEMBLANCE	semblance-identity	54.55%	6
	semblance-similarity	45.45%	5
COMPARISON-SPECIFICATION	general-comparison	51.04%	49
	experientialised	48.96%	47
EXPERIENTIALISED	--	100.00%	47
	purposive	0.00%	0
EXPERIENTIALISED-1	comparison-quantity	44.68%	21
	comparison-quality	55.32%	26

Table 2.5 System of Participant Identification: Descriptive results for combined texts.

Within Presuming reference, these texts reveal interesting marked examples of Unique reference, in comparison to the unmarked realisation as a proper noun. Unique reference is used for a label referring to a diagram – when the diagram is 'named' in the text – or another part of the book or chapter, and for when quotation marks are used (e.g. "ground loops" in AN). The use of quotation marks here is hypothesised to represent a Unique term that should be known to the discourse community, since the author does not fulfil a prediction of advanced labelling (Tadros, 1989, see 4.1) by elaborating on this term. Examples of unmarked Unique reference for 'proper nouns' include IEEE (BN), AM and FM (AMFM), and UNIX and MS-DOS (FSI). Also in FSI the terms CREATE and READ are treated as Unique. Unique reference totalled 32 of 1256, or 2.5%, Presumed participants.

Of the remaining 1224 Variable groups, only 156 (12.75%) are pronominal. These are divided between 86.5% non-interlocutors and 13.5 % interlocutors. All of the instances of Interlocutor occur in three different sources: Buchla and McLachlan (1992) Coulouris et al. (2001) and Horowitz and Hill (1989). It would appear, then, that the choice to address the reader or to refer to the authors directly is a consistent choice by each author.

Only one text (AN) does not include the category of Directed. The 50 examples of *this*, *that*, *these* and *those* are spread across the other texts, with RM having a disproportionately low number of realisations. Just one of the 1018 Undirected groups include Questioning mood. The majority of Undirected groups are unmarked, with only six percent (52) forming a Superset, divided into 55% non-selective and 45% selective. The results for the different terminal options are: in selective, 18 occurrences of Quality (e.g. *the outer end, best, highest, primary, lowest most common*), in non-selective, 21 occurrences of Individuated (e.g. all using *each*), and 11 of dual - Inclusive (*both*) and 2 of dual – Alternative, both positive (*either).*

In 2305 instances of Presenting and Presuming reference, only 159 groups (6.9%) are pronominal. It could be deduced from this figure, and the preceding paragraph, that there is a very high proportion of nouns to pronouns in these texts compared to other genres. However, it is difficult to make such comparisons without valid data from other genres. The sample narratives provided in Martin (1992) suggest that a higher proportion of pronouns would be expected than the approximately 1 in 14 participants found in these texts. This may be the result of a lot of new participants being introduced, and a correspondingly higher proportion of Presenting to Presuming reference than may be

expected in, for example, narrative. A higher proportion of generic reference may contribute to the increase in the ratio of Presuming to Presenting reference, but as a result of the highly context-dependent nature of Generic reference, this analysis tends to select Specific whenever there may be doubt. It could be predicted that such a low ratio of pronominalisations may lead to a rather 'heavy cognitive load', as there seem to be few participants becoming a 'focus' in the text, making it appear rather dense. This observation is supported by the use of only one Directed Presuming nominal group. Until the behaviour of these participants is tracked, however, it would be very difficult to make reliable comments on how these proportions affect the texture of these texts (see Table 2.5).

The choices in the Presenting system are typically Partial (98.3%). All but 1 of the examples of Total – Nominal and Pronominal – use the determiner *all*. The one exception is *everyone else*. Not surprisingly, 82.6% of all Presenting options are Unmarked, making 90.7% of the Partial-Nominal category. Typical examples include *a physical phenomenon*, *as a test of, systematic errors, access, different lengths of*, and *a separate built-in motor*. The marked option of Major Role is used in only one text, FMAM, with 3 instances of *certain* and 1 of *so-called*. In all, 89 choices of Presenting reference are marked. They are split into 9 instances of *any* in 4 texts as Unrestricted, 38 Particular, realised by *some*, *most*, *particular*, and *typical* amongst others, and 38 Non-particular across all texts, exemplified by no use of determiner in the group *wide area, open distributed systems* in Sec.

The lack of Major Role in the majority of texts (Table 2.5) may indicate that the authors do not wish to accord special status to one of the Participants in this classification text. Within the hierarchy of knowledge that is being described here, all participants are equally important in the structure of knowledge. Consequently, when a Participant is Presented in a marked way, it will not be accorded higher status, but may need to be separated from other members of the group for textual reasons. The example given for Non-Particular would support this view.

Only 96 (4.2%) of the Specific groups enter the Comparison system simultaneously with the Presenting/Presuming system and Specific/Generic. Of these the 85 comparisons of difference come from all texts. Typical realisations include *in the latter case*, *less bandwidth*, *over much further distances*, *a separate built-in motor*, *the upper cable* and *Alternative hand movement*. The 11 Semblance comparisons are split between 6 of the Identity type, realised by *the same*, *of each*, and *the latter,* and 5 of the Similarity type, realised by *such a*, *similar*, and *other*.

Simultaneous with Comparison-Similarity choices are Comparison-Specification. The majority of these (80%) are General-Comparison. Typical examples are *in the same way as* and *the different kinds of*. Neither of the Experientialised comparisons are of the purposive type. One is Experientialised-Quality (*the outer end*) and one Experientialised-Quantity (*different lengths of*).

Specific or Generic is selected 2112 (91.6%) and 193 (8.4%) times, respectively. The relatively high proportion of Generic compared to Specific is to be expected of these texts. An alternative analysis may have reached a higher figure for Generic, but bearing in mind the pedagogical Register of these texts I have attempted to show whenever possible that a knowledge of the context of culture is not always necessary.

2.6.1.2 Issues and Problems in Analysis

This section describes some of the more difficult choices in text analysis, with examples, in the analysis of Participant Identification. In general, it must be pointed out that the analysis of Participant Identification presented far fewer issues, and raised fewer questionable or borderline decisions than the choices for Tracking (see 7.2 below). That is to say, the system of Participant Identification proved robust when applied to these texts.

Headings in these texts are typically analysed as Neutralised in the Identification system. This allows a participant to enter the text for Tracking, without having being designated as Presented or Presumed. Because the heading is marked as such, it allows a reader to enter the text at that point, without having to retrieve the Identity of the participant from a previous point in the text. This gives us our first piece of evidence that the Identification and Tracking networks are independent, because while the Identification system is neutralised, the Tracking system can begin to operate.

In many cases, it is almost impossible to distinguish between Generic and Specific, such as in the first sentence before a co-text has been developed. For instance, the group *the desired signal* in the sentence 'the ultimate limit of detectability of weak signals is set by noise - unwanted signals that obscure the desired signal' (AN) could be analysed as Generic, since it can be identified as any desired signal in any context. In this instance it was analysed as specific as a result of its standing in contrast to *unwanted signals*. The point is that both analyses are possible. In this and most instances I chose to allow the weight of

the co-text to guide the selection, and wherever possible allowed the text to provide specific reference.

2.6.2 Participant Tracking

The system of Participant Tracking enables the reader to follow participants through a text. This is achieved through the simultaneous systems of grammatical phoricity and lexical relations, which allow the reader to follow the same participants through different processes in a text. Typically written text depends largely on the context of situation, particularly the verbal co-text, to track participants. The discussion above (see section 2.4.2.4 to 2.4.2.4.4) hypothesises that once a relationship has been established through grammatical phoricity, lexical relations are ordered roughly in terms of markedness; at each point of selection in the network, the upper choice in Fig. 2.13 is hypothesised to be more frequent and salient than the lower choice(s).

2.6.2.1 Typical Analyses and Realisations

Fig. 2.13 illustrates the likelihood of each choice in the network, according to the results of these texts. The number and percentage of features in the identification system are detailed in Table 2.6. The following is a summary, with comments, of these results.

Of the 2383 participants in the analysed corpus, the authors indicate that the reader is not expected to track the identity of 881 (37%) of them from anywhere else – that they are an Addition at that point (Table 2.6). Typical examples include *a physical phenomenon* (EAP), *radiofrequency interference* (AN), *an entire file* (FSI), and *precision bearings* (MM) and *numerous benefits* (FMAM). It is likely that a ratio of 2 trackable participants to every 1 new participant is lower than in other genres, or at least that the taxonomic nature of these texts has an effect on these figures.

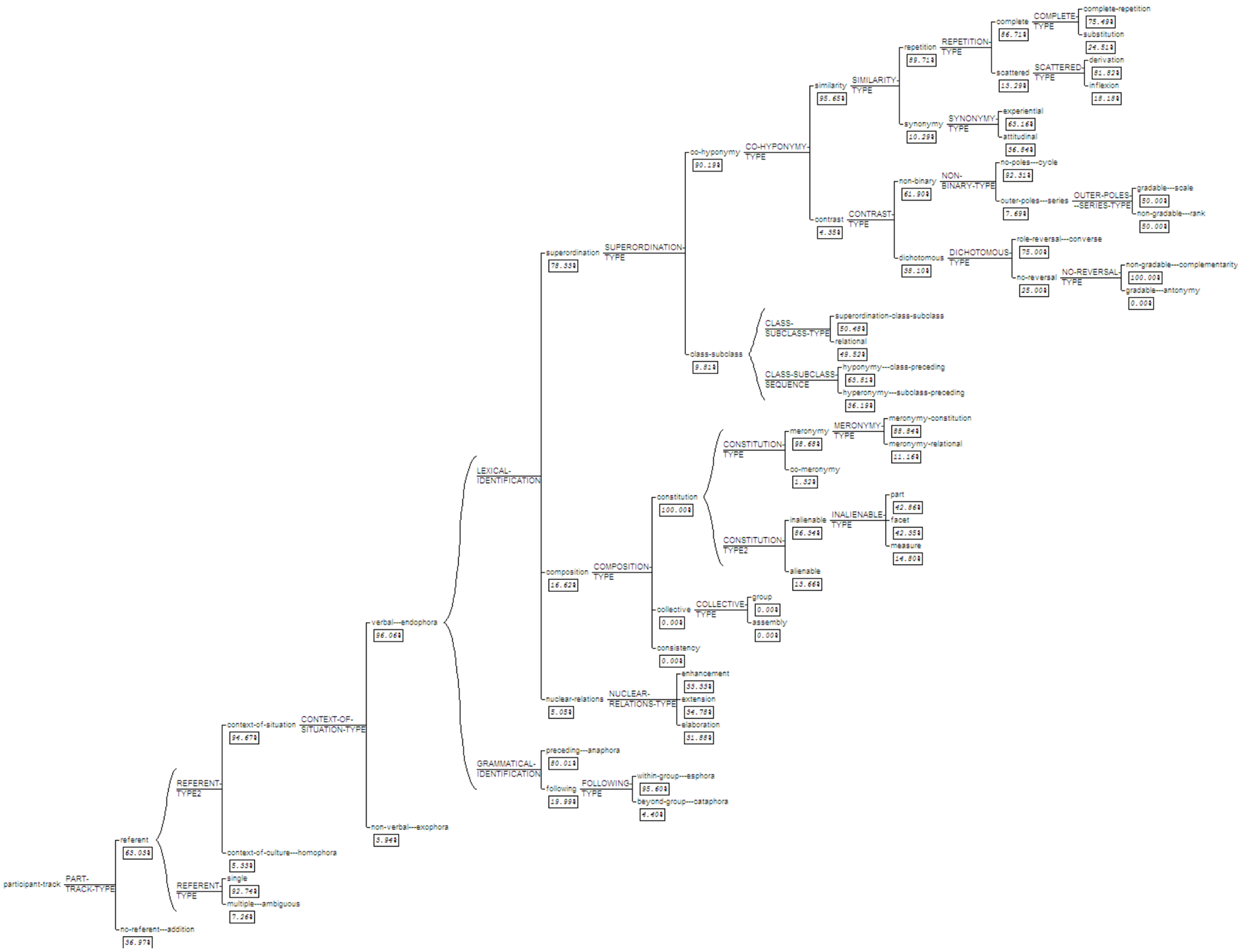


Fig. 2.13 **Selection of Options, as Percentage, of Participant Tracking Network for Combined Texts**

Feature		Mean	N
PART-TRACK	referent	63.03%	1502
	no-referent---addition	36.97%	881
REFERENT-TYPE2	context-of-situation	94.67%	1422
	context-of-culture---homophora	5.33%	80
CONTEXT-OF-SITUATION	verbal---endophora	96.06%	1366
	non-verbal---exophora	3.94%	56
LEXICAL-IDENTIFICATION	superordination	78.33%	1070
	composition	16.62%	227
	nuclear-relations	5.05%	69
SUPERORDINATION	co-hyponymy	90.19%	965
	class-subclass	9.81%	105
CO-HYPONYMY	similarity	95.65%	923
	contrast	4.35%	42
SIMILARITY	repetition	89.71%	828
	synonymy	10.29%	95
REPETITION	complete	86.71%	718
	scattered	13.29%	110
COMPLETE	complete-repetition	75.49%	542
	substitution	24.51%	176
SCATTERED	derivation	81.82%	90
	inflexion	18.18%	20
SYNONYMY	experiential	63.16%	60
	attitudinal	36.84%	35
CONTRAST	non-binary	61.90%	26
	dichotomous	38.10%	16
NON-BINARY	no-poles---cycle	92.31%	24
	outer-poles---series	7.69%	2
OUTER-POLES---SERIES	gradable---scale	50.00%	1
	non-gradable---rank	50.00%	1
DICHOTOMOUS	role-reversal---converse	75.00%	12
	no-reversal	25.00%	4
NO-REVERSAL	non-gradable---complementarity	100.00%	4
	gradable---antonymy	0.00%	0
CLASS-SUBCLASS	superordination-class-subclass	50.48%	53
	relational	49.52%	52
CLASS-SUBCLASS-SEQUENCE	hyponymy---class-preceding	63.81%	67
	hyperonymy---subclass-preceding	36.19%	38
COMPOSITION	constitution	100.00%	227
	collective	0.00%	0
	consistency	0.00%	0
CONSTITUTION	meronymy	98.68%	224
	co-meronymy	1.32%	3
MERONYMY	meronymy-constitution	88.84%	199
	meronymy-relational	11.16%	25
CONSTITUTION-TYPE2	inalienable	86.34%	196
	alienable	13.66%	31
INALIENABLE	part	42.86%	84
	facet	42.35%	83
	measure	14.80%	29
NUCLEAR-RELATIONS	enhancement	33.33%	23
	extension	34.78%	24
	elaboration	31.88%	22
GRAMMATICAL-IDENTIFICATION	preceding---anaphora	80.01%	1093
	following	19.99%	273
FOLLOWING	within-group---esphora	95.60%	261
	beyond-group---cataphora	4.40%	12
REFERENT	single	92.74%	1393
	multiple---ambiguous	7.26%	109

Table 2.6 System of Participant Tracking: Descriptive results for combined texts.

2.6.2.1.1 Phoric Relations

A relatively low proportion (5.33%) of the 1502 participants that can be tracked to another referent depend on homophoric reference, rather than on the context of situation, to track their identity, and these tend to be distributed unevenly across the different texts. This is typical of technical text more than a more informal style because the technical text will attempt to create as much of its own context as possible while informal text will attempt to use the solidarity between interlocutors to depend on a presumed shared context. Three texts (EAP, RM and FSI) contribute 51 of the 80 instances, and there are no instances of homophora in MM, AN or Arc. In the RM text, for instance, *the ground crew* in the sentence "On aircraft fitted with mechanically or electrically operated hooks, the hook usually has to be raised manually by the ground crew using a lifting rig or mechanical hoist" is presumed from the context of culture as it uses presuming reference, but, unlike other participants with presuming reference, has no place in the co-text that it can be related to. That is, for this writer it is 'common sense' that every airport has a ground crew and so you naturally presume that they will be there to raise the hook manually.

Within the context of situation, 3.9% of participants (56) are tracked to non-verbal participants (see Table 2.6). The main types of reference here include referring directly to the reader (*you* in EAP) and writer (*we* in Arc, Sec and AN), and referring to other parts of the text (*section* in BN, *chapter* in Sec) or to non-verbal aspects of the publication (*fig.* In RM, MM and FSI, and *In the model* in MM). If we consider an illustration as a non-verbal part of the context of situation, then the majority of exophoric participants in MM are tracked to illustrations accompanying the verbal text. They are tracked both explicitly, as in *Fig. 11.5* and in *the model (shown)*, and implicitly as in *the cutting-fluid reservoir*, and in the title *Overarm and arbor support*, which are two labelled parts in an earlier diagram. The precise way that the textual metafunction enables these cross-modal references to multiply meanings (Lemke, 1998) needs to be investigated in further research.

Simultaneous with the choice of Context of Situation or Context of Culture is the possibility of multiple referents. It is quite possible that the score of 7.3% (109 instances) of Multiple-Ambiguous reference is an underestimate. This is largely because it is natural and often very simple, as a reader, to identify the most obvious single referent that is being presumed in a cohesive tie. Frequently, however, it is possible to identify other candidates that are in the co-text, that perhaps have been previously pronominalised, and

that match grammatically with the participant being analysed. The following example of *This* is analysed as ambiguous because *applicational error* could refer either to the specific example (using a voltmeter) or the general case (misuse of an instrument):

> 2.1 Sometimes a systematic error occurs because of the misuse of an instrument outside its design range, such as when a voltmeter is used to measure a frequency beyond its specifications. (*This* is also called an applicational error.) (EAP, p.39)

All texts exhibited examples of ambiguous reference, with EAP and FMAM containing the highest counts of 47 and 29, respectively.

The majority of groups that can be tracked (anaphorically, esphorically or cataphorically) to another participant (96%) have an endophoric, or verbal, relationship. This represents 57.3% of all participants. This high proportion is to be expected in written technical text as the text itself is the main source of meaning-making. Also typical of technical text is the relatively high proportion of Esphoric reference in these texts. As per Martin's (1992) prediction, Esphoric (within-group) reference is more frequent than Cataphoric (across group) reference by a ratio of about 19:1 (95.6%:4.4%). Anaphoric relations outnumber Esphoric and Cataphoric reference by a ratio of about 4:1 (80%:20%).

Significantly for this study (see discussion in section 2.4.2.4), each endophoric referent exhibits not only a grammatical relationship, but also at least one identifiable lexical relationship with another referent. In fact, in many instances it was possible to posit multiple lexical links. The lexical features of endophoric reference are exemplified in the following section.

2.6.2.1.2 Lexical Relations

Looking in detail at the lexical relationships, we see that Superordination accounted for 78.3% of endophoric reference, relations of Composition for 16.6%, and Nuclear Relations for the remaining 5.1%. While this last figure may appear low, it must be remembered that the decision to place the lexical relations in a presumed hierarchy, and to select only the highest option, is bound to affect these results. That is, in analysing the text it is clear that the lexical nature of many of the phoric relations could be tracked in a number of ways. It is also clear that the higher up the hierarchy the relationship, the easier it is to justify the relationship. The figures for both Nuclear Relations and Composition, therefore, exclude a relation of Superordination, and the figures for Nuclear Relations necessarily exclude relations of Composition, while the reverse cannot be certain. The discussion below proceeds from the least to the most frequent systems.

Nuclear Relations are divided very evenly among relations of Elaboration (31.9%), Enhancement (33.3%) and Extension (34.8%) (see Table 2.6). The lexical relation of Elaboration is exemplified by *the means* in "The two keys again provide the means of transmitting the drive" (MM). In the sentence "Clamping of the saddle to the knee is achieved by two clamps on the side of the saddle" (from MM), the participant *Clamping* is analysed as having an esphoric (group-internal) relationship with the rest of the group, with the participant providing Enhancement within the group. An example of Extension is *operation* in the following extract:

> 2.2 *Wheel brakes* form the primary method of retarding the forward movement of most aircraft when on the ground and, in common with most braking systems, *they* rely on the principle of energy conversion for *their operation*. (RM, section 4)

The example here is analysed as Extension because of the relationship of *operation* with *wheel brakes* (and *braking systems*), to which *operation* is tied by *their*. The semantic relationship between these two items – the relationship that enables the tie to be made – is not one of Mernoymy or Superordination, for example, but one of Process + Range, as in Table 2.4, even though *operation* is a grammatical metaphor for the congruent verbal process.

Relations of Composition account for only 16.6% of lexical relations. Within Composition, all 227 groups exhibit a relationship of Constitution. This is likely to be a result of the type of texts being analysed: all texts describe the component parts of a larger system, although clearly they are not necessarily confined to that genre. More surprising, perhaps, is that only three Constitution relations (1.3%) are analysed as Co-meronymy. That is, these texts tend to describe systems hierarchically rather than relating arrays of sub-components to their equivalencies. Within Meronymy, the unmarked relationship of Meronymy (88.8%) outnumbers Meronymy-Relational (11.2%) by a ratio of about 9:1. Typical examples of Meronymy include groups such as *bottom end*, *its top surface* (MM), *the resolution of*, *the correctness of* (EAP) and *for its employees* (Sec), while Meronymy-Relational is typically realised by items such as *structure, distribution, direction* and *value*. 196 instances of Inalienable meronymic relations contrast with 31 for Alienable. These relations are realised by *quality*, *precision*, *correctness*, and *features* (Inalienable) and *privacy*, *security*, *aspect*, and *problem* (Alienable) amongst others. Inalienable meronymy is further divided into 42.4% Facet, 42.8% Part and 14.8% Measure.

The third main category – Superordination – contains the majority of examples of lexical relations. As noted in the discussion above, these are the majority partly because they are the first to be selected. That is, it is often possible to assign more than one lexical or semantic characteristic to a phoric relationship, but this model posits a hierarchy of lexical relations based on presumed ease of processing, as set out in the network diagram (appendix 2.2). Consequently, Superordinate relations can be predicted to be more frequent than the others. A relation of Superordination may be the only tracking lexical relation for a group, or it may subsume other relations.

Of the 1366 Endophoric relations that enter the simultaneous lexical and grammatical tracking relations, 1070 (78.3%) were judged to show a relation of Superordination, (producing a ratio of approximately 31:6:2 for Superordination: Composition: Nuclear Relations). According to the network, the easiest, or most salient, lexical relationship to process is that of Complete-Repetition, *i.e.* a participant is repeated exactly. In this case we find that 542 (50.7%) of the groups with a relation of Superordination terminate in Complete-Repetition. To give an indication of the prevalence of this relation, we see that 542 groups represent 22.7%, or approximately two in nine groups, of the whole text.

The alternative to Complete-Repetition is Substitution. Typically, a participant is substituted by a co-referential lexical item, such as a pronoun, but Substitution here allows lexical items that are Co-hyponymous but not the same lexical item repeated, as in Complete Repetition. That is, lexical items with a reminding function that are co-referential can be used to track the same Participant. For example, if a participant has been modified in some way, but is co-referential, it will be analysed as Substitution. There are 176 realisations of this choice, which reveals a relationship of 3:1 in favour of Complete Repetition. As well as common pronouns, the range of realisations includes *the same*, *each*, *both*, *the latter* and *only small pieces*. A further level of delicacy could be added to the network as a response to substitution items which name the relationship. For example, since items such as *term* can be used to both substitute for another term and name that relationship, the label used in other parts of the network, Relational, could also be placed here.

Combined, Complete-Repetition and Repetition-Substitution constitute 30.1% of all participants, and 52.6% of all phoric groups across the two texts. Fig. 2.13 shows that these texts follow the predicted hierarchy for the most common choices in Superordination, with a proportion of approximately nine to one in favour of the choice

hypothesised to be the easier to access cognitively. This proportion mirrors examples identified by Halliday and James (1993) when looking at other grammatical systems. While the less likely choice is highly marked and salient, in terms of Information Theory (Shannon, 1948; 1950/1993), the unmarked choice adds little information to the clause.

The next set of Repetition choices include Scattered-Inflexion (20 realisations) and Scattered-Derivation (90 realisations). A typical example in MM is the group *transmitting* in the group *the means of transmitting the drive*, derived from *transmit* in an earlier clause. One fairly extreme example of Derivation is *the amplification* (AN), as the only previous mention of this lexeme is in *Amplifier* in the title. Contrasting with Repetition are the 60 realisations of Synonymy-Experiential and 35 of Synonymy-Attitudinal.

In Co-hyponymous relations, the choice is between Similarity, covering the Superordination choices discussed so far, and the next set of relations covering Contrast. Within the whole text, Similarity relations account for 38.7% of all participants while Contrast relations are only 1.8% of the text. Contrast relations are fairly evenly divided between 26 realisations of Non-binary and 16 of dichotomous. Table 2.7 summarises Contrast relations (with examples from MM).

System	Choice	#	% in choice	Milling Machine (MM)
NON-BINARY-TYPE	no-poles-cycle	24	92.3	*6 to 250 mm/min*
	outer-poles-series	2	7.7	-
DICHOTOMOUS-TYPE	role-reversal-converse	12	75	*Each direction*
	no-reversal	4	25	*(see below...)*
NO-REVERSAL-TYPE	Gradable-antonymy	0	0	-
	non-gradable-complementarity	4	100	*Alternative hand feed*

Table 2.7 Summary and Realisations of Contrast Lexical Relations

In the Superordination network, Co-hyponymy relations contrast with Class-Subclass relations. In these texts the proportions are Co-hyponymy 965 (90.2%): Class-Subclass 105 (9.8%), or approximately 9:1, lending further support to the hierarchy of semantic relations posited in this study. In the first simultaneous system, the Superordination-Class-Sub-Class type is selected 50.5% of the time while the Superordination-Relational type is selected 49.5% of the time, a ratio of about 1:1. The other simultaneous system Class-Subclass Sequence reveals a division between Hyponymy (class-preceding) with 63.8% of realisations and Hyperonymy (subclass-preceding) with 36.2%, a ration of approximately

2:1. These ratios represent a clearly marked relationship (9:1), where the marked partner is unlikely to occur so it has much greater saliency, a slightly marked relationship (2:1), and an unmarked relationship (1:1) where the equal probability of the two choices allows little opportunity to make meaning.

2.6.2.2 Issues and Problems in Analysis

Initially, this section will deal with specific problems in analysis of the texts, and then discuss more general issues of this aspect of the textual analysis.

Before we look at examples of the tracking of participants we need to look at non-participants, in particular those examples that may contain a nominal group but fail to qualify as a participant – an agent or medium in an ergativity analysis. As predicted in the discussion on the nature of participants (see sections 2.1 and 2.4), some groups that fulfil the same syntactic role as a Participant were not identified as such:

> The word there in such clauses is neither a participant nor a circumstance – it has no representational function in the transitivity structure of the clause (Halliday and Matthiessen, 2004 p.257)

In extract 2.3, *there* substitutes for a participant in terms of clausal position but does not enter an ergative analysis as either agent or medium. It is therefore not considered a participant (but will play a significant role in Theme analyses in the following chapter):

2.3 On some aircraft *there* is also a mechanical indicator on the release unit which normally lies flush with the aircraft structure (RM, section 25)

In the sentence "The saddle has dovetail guideways on its upper surface, at right angles to the knee guideways," the group *at right angles* is considered neither a participant nor a potential participant because it is a fixed phrase that cannot act as agent or medium in the clause even though *right angles* has the form of a nominal group. Some common lexical phrases were included as participants under the identification of Generalized reference, including *by chance*, *at all times*, and *for example*. Since these could then be tracked through various lexical relations (*by design*, *at other times*, and *in this example*, respectively) they were analysed as participants, even though it would be difficult to analyse them as anything more than potential participants. That is, the final decision on whether to include a potential participant was based on the ability to track it later in the text.

Section 2.4.2.4 discussed the reasons for only analysing lexical relations when there is a grammatical relation already indicated. We can now look at an example. In the sentence 'The front of the column carries the guideways upon which the knee is located and guided in a vertical direction', it is possible to identify a range of lexical relations for the group *a vertical direction*. There are certainly Nuclear Relations with *located* and *guided*, and possible other semantic relations to *guideways*. Further, *a vertical direction* could also be analysed as repetition from an earlier sentence. However, it is assumed that since the reader is not required to make the connection, as there is no grammatical reference, the lexical relations may go unanalysed. This is one example where it is possible to identify lexical relations, but since no grammatical relation is realised in the text, the lexical relations are not analysed. That is, the reader is not required to connect the lexical items with each other, as there is no grammatical indication that the items are related. Consequently, while it may be possible to analyse lexical relations, these relations will not be assumed unless they are simultaneously indicated grammatically.

There appears to be a 'gap' between choice of Complete-Repetition and Substitution. That is, Complete-Repetition should include only the exact repetition of a lexical item. Substitution, then, must cover all forms of replacement that do not depend on a change by Inflexion or Derivation. The discussion below deals with this issue in detail. In the first example of:

> 2.4 Cutters can be mounted straight on the spindle nose or in cutterholding devices which in turn are mounted in *the spindle*, held in position by a drawbolt passing through *the hollow spindle* (MM, p.185)

the group *the hollow spindle* is not an exact repetition of the previous co-referential group *the spindle*. Further modification has been attributed to the same group, and so the relationship of Similarity holds, but not by Derivation or by Inflexion since the word class has not been altered. In this study, I have favoured Substitution, but it still is not as frequent as Complete-Repetition. By default, then, the group is analysed with a relation of Substitution – a different form representing an identical co-referential group. This is not a typical example of pronominalisation, but could be said to perform the same function. The next few examples deal with this issue further.

The example *a longitudinal direction* (MM) was analysed as Experientially Synonymous with *transverse direction*, *a vertical direction*, and *transverse movement*. This does not mean that *a longitudinal direction* is directly equivalent in meaning to these other terms. Instead, these terms, and other participants (*a vertical direction*, *guideways*, *vertical movement*, *transverse movement*,

Drive, *guiding*, *transverse direction*) form a distinct lexical set, with members exhibiting derivational and experiential variation. It is also be possible that, through derivation, Participants can be tracked to non-participants, as grammatical metaphor enables a range of potential participants to become participants (section 2.1). The term Synonymy may be confusing as it suggests a semantic equivalence. Synonymy is not intended here to represent identity or co-reference. A participant may be tracked by its semantic relation, characterized by the Tracking network, to another participant but the participants do not have to be co-referential. To take a quick example, the relationships of Converse and Complementarity (see Fig. 2.13) define both partners in a dichotomous relationship. If 'the winner' is mentioned, its Converse 'the loser' can be tracked through the first partner, as the 'frame' (Emmott, 1994) contains both partners. Consequently, the term Similonymy (see section 2.4.2.4.1) is more appropriate as it implies a relationship of similarity rather than identity.

One particularly problematic example is *the drive* in the sentence 'The spindle, accurately mounted in precision bearings, provides the drive for the milling cutters.' The analysis here selects the unmarked Esphoric relationship with *the milling cutters*. However, the type of lexical relationship between *drive* and *milling cutters* is not immediately apparent, although a type of grammatical metaphor has been employed here, corroborated by the more apparent Nuclear Relations between *provide* and *the drive*. This is problematic because *the drive* could also be analysed as having either a lexical relationship of Co-meronymy, or a Nuclear relation of Elaboration, with *The spindle*, again suggested by the grammatical metaphor. Where there is a grammatical metaphor it is helpful to code an item for both its metaphorical meaning and its congruent meaning. The restrictions placed on the analysis, as a result of the software being used, however, are made most apparent here, as the computerised tagging of the group accepts just one analysis.

In another example of this methodological restriction, in the sentence 'The spindle runs in a quill which is moved through a rack and pinion in the same way as a drilling machine spindle', the group *in the same way* is analysed as Complete Repetition-Substitution. However, it could also be analysed as not being a participant, since it is a fixed phrase, or even as a Nuclear Relation of Elaboration. Only the first of these analyses has been recorded. As mentioned previously, all lexical relations of Superordination may include within them lexical relations of Composition and/or Nuclear Relations. Composition relations may also hide results for Nuclear Relations, because in each case only one coding

was accepted for each group. Only Nuclear Relations can be safely predicted to include only that type of lexical relationship.

A number of problematic analyses centred on Nuclear Relations, largely because the participant analysis tended to select other participants to identify lexical relations. The phrase *the presence* in the sentence '*Even if the quantity being measured is not weak, the presence of noise degrades the accuracy of the measurement*' (AN) provides an interesting example of Nuclear Relations. In contributing very little in Ideational terms to the nominal group or clause, this phrase resembles a 'delexical verb' (Sinclair *et al.*, 1990). It also exhibits a Nuclear Relation of Elaboration to the rest of the group; the textual reasons for including *the presence* in the nominal group seemingly outweigh the ideational reasons for not including it. Including the element *presence* in the group does change the sequencing of items, and may alter the meaning of the clause and for this reason was included in the analysis. In some other cases Nuclear Relations produced unclear analyses.

One area of cohesion studies is related to ellipsis. Continuing with the same example, the group *the quantity* can be considered to be elliptical for *the quantity of noise*. In the analysis here, there are no considerations of what may possibly have been meant, only what has been written (or 'worded'). For participant tracking analysing ellipsis could possibly lead to difficulties as a participant may be tracked through two Ellipses to another Substitution or Complete Repetition, but it would not be counted as such. Further research could be directed towards the investigation of this issue.

2.7 Directions for further Development

The study in this chapter has attempted to demonstrate that the SFL approach to Participants, derived from Martin (1992), offers a reliable comprehensive model for analysing discourse. A number of alternative approaches have been rejected in favour of the SFL model, but, as pointed out in the results sections, there are still a number of areas where SFL can be improved. For instance, currently there is no model to combine the results of these two systems (see appendix 2.5 for quantitative example). Further issues that would improve the model are discussed below.

2.7.1 Computational Studies in Centering

The model in this study does not specify, but also does not preclude, Prospection. It may be possible to include a more dynamic approach to Participant Tracking (or Continuity) that would allow for a greater role for Prospection, perhaps along the lines of a Discourse History, as outlined in Matthiessen (1992). This produces a list of participants – similar to Centers (Grosz, Joshi and Weinstein, 1995) or the s-list of Strube and Hahn (1999) (see below) – that are presumed to be potential foci at any point in discourse. This study does not specify a dynamic list such as this; neither does it specify a retrospective view of anaphora that would prohibit Centering.

The main contribution to cohesion claimed by Centering studies is that, while processing text, readers do not look back to track participants, as suggested in Halliday and Hasan (1976); rather, as each participant is introduced it is *carried forward* and marked for potential syntactic and functional roles. Emmott's (1994, 1997) proposal that prospection plays at least as important a role in anaphora resolution as retrospection is supported by the Centering approach.

When applied to anaphora resolution, Centering (e.g. Grosz, Joshi and Weinstein, 1995; Grosz and Sidner, 1998; Tetreault, 2001) generates a series of conditions, prepared as algorithms, that dynamically determine which participants are most likely to be taken up as the focus of attention in the following discourse. Significantly for this study, Strube and Hahn (1999; Strube, 1998; Navarretta, 2002) introduce more 'functional' elements, including information structure (mainly realised by sequence), into their model of centering to produce higher levels of resolution than the statistically-based studies reported in 2.3.

Using Grosz's (1977; 1978) notion of focus, Sidner identified how pronouns (1981) and definite noun phrases (1978) can co-specify an antecedent noun phrase in discourse. That is, the discourse focus is most likely to be the element that becomes an anaphor in later discourse, if it is not already. Following Grosz, Sidner distinguishes between a discourse focus and an actor focus. The discourse focus is defined as "a process of tracking the speaker's foci as they change over the discourse" (Sidner, 1981, p.220) while the actor focus "is a discourse item that is predicated as the agent in some event" (p.220). Grosz and Sidner (1998) describe how notions such as discourse focus and potential focus are directly related to the backward-looking center (C_b) and forward-looking centers (C_f),

respectively. The seminal Centering paper (Grosz, Joshi and Weinstein, 1995) made the important link between the centers and linguistic form.

Fig. 2.14 illustrates the prospective nature of Centering theory. For any one sentence, Centering Theory posits one, and only one, *backward-looking* center (C_b) selected from a set of *forward-looking* centers (C_f) from the previous sentence. The C_b is the main focus of that sentence. The C_f is the set of possible discourse foci of the following sentence and is identified syntactically according to the following hierarchy: "SUBJECT > OBJECT(S) > OTHER" (Grosz, Joshi and Weinstein 1995 p.214). The set C_f is dynamically moved forward to the next sentence. Consequently, each sentence (Utterance or U) develops a list of C_f for the following sentence (U_{+1}). The C_b in U+1 will be the highest ranking item which grammatically agrees for gender, number and role from the set of C_f in the preceding sentence (U). As pointed out by Strube and Hahn (1999), Centering Theory is fundamentally syntactic – the Center is not chosen by the speaker, but is defined by syntax.

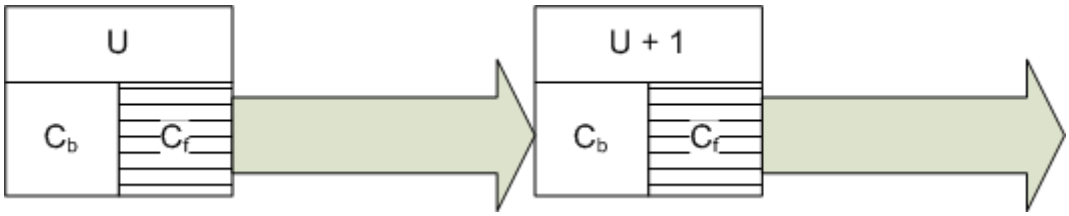


Fig. 2.14 **Centering rules: Select one Backward-looking Center from the Previous Set of Forward-looking Centers**

2.7.2 Developments in Centering Theory

Further research to extend Centering theory includes studies that directly address the claims proposed in Grosz, Joshi and Weinstein (1995), while others take a more critical stance and try to establish the validity of the underlying concepts. Perhaps the most significant development in Centering Theory in relation to this study is the work by Strube and Hahn (1999; Hahn and Strube, 1997; Strube, 1998). In their model, an S-list is generated instead of a set of C_f, based not on syntactic status but on functional criteria. Here they develop the claim that syntactic, pragmatic and semantic factors combine to determine Centering.

Strube and Hahn draw heavily on Prince's (1981) taxonomy of Given and New to order participants in terms of salience at any discourse moment. Fig. 2.15, taken from Strube

and Hahn (1999), divides discourse entities into hearer-old (OLD), mediated discourse entities (MED) and hearer-new (NEW). OLD is divided into Evoked (E) and Unused (U), corresponding to discourse-old and hearer-old, respectively (Prince, 1981). Within Mediated (MED), the three categories are Inferable (I), Containing Inferable (I^C) and Anchored Brand-New (BN^A). All MED terms are derived from Prince's view of Bridging. Finally, NEW contains only one category: Brand-New (BN). These categories are then ordered to predict the likelihood of co-specifying a pronominalised element in the subsequent sentence. In Fig. 2.15 the order is from left to right: "the extended C_f ranking … will prefer OLD discourse entities over MEDiated ones, and MEDiated ones will be preferred over NEW ones." (Strube and Hahn 1999 p.323). Within each category, the different sub-types are not hypothesised to show any preference or ordering.

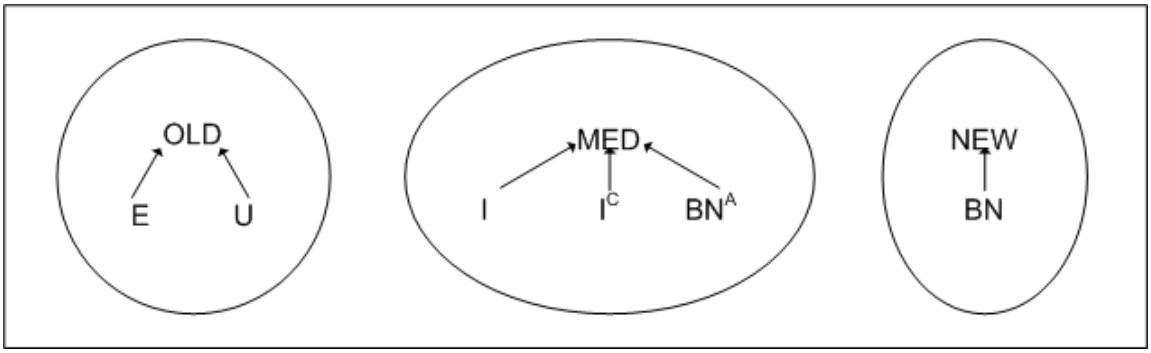


Fig. 2.15 **Information Status and Familiarity (Refined Version) (from Strube and Hahn 1999 p.324)**

As with other computational studies of anaphora resolution, a major criticism is that, for the sake of accuracy, a wide range of linguistic features are ignored in analysis. Strube and Hahn examine only 3rd person personal and possessive pronouns, and what they call 'functional anaphora'. That is, they do not evaluate the efficacy of their method of anaphora resolution on the full range of personal pronouns, they do not track the antecedents of definite noun phrases, and they do not examine all of the possible realisations of participants. For this narrow range of features, Strube and Hahn claim much greater efficacy in pronoun resolution than Grosz, Joshi and Weinstein (1995), and so can claim to have more efficient factors that contribute to the set of C_f. However, as pointed out in section 2.2 and 2.4, the claims made by theories based on the concept of 'Bridging' must be evaluated carefully in the light of a critical discussion of the methods and conclusions reached by Clark and colleagues. Similarly, Prince's taxonomy (see section 2.2.3 for discussion) has not been subjected to critical review or empirical evaluation.

One suggestion from Centering studies that does stand up to closer examination is the concept that centers, or participants, are *carried forward* in discourse; that as each participant is introduced it *prospects* to its likely pronominalisation and its focus. This proposal is examined further in section 2.7.3.

2.7.3 Prospection in Text

Models of cohesion that depend on a retrospective view of anaphora, and other reference phenomena, do not reflect the processes employed in discourse, according to preliminary psychological studies. The limited understanding that we now have of the psychological processes involved demand a prospective, dynamic model of language, including reference or participant tracking, alongside a retrospective static one. That is, although it is possible to analyse pronouns as 'referring back' in text, it is also possible that participants are carried forward in text ready to pick up the referential devices that follow them. To establish the validity of prospection, it is examined briefly as explicit prospection at the broader level of discourse, before looking at the prospection of participants in more detail.

Networks in Systemic Functional Linguistics can be seen from the two perspectives of process and product, of dynamic and synoptic (see section 1.2.2). The final choice in a network can be seen as the product of different choices in the network. Alternatively, the choices present at each point in the model plot a particular route taken through a network that may reflect the processing of meaning. Further, as a text proceeds, Participants are introduced, change and are dropped from focus. Participant Tracking and Participant Identification, by their very nature, are dynamic processes. Many models of cohesion accentuate a retrospective, static view of reference in text. While this is one part of an analysis, the dynamic nature of participants must also be examined. This section outlines other perspectives that may contribute to a dynamic approach to participant tracking.

2.7.4 Explicit Prospection

In some cases, text is explicitly prospective, making clear what general points or generic moves (Swales, 1990) the reader can expect. Tadros (1985; 1989) describes the various ways that a text makes predictions, including Enumeration, Advance Labelling, Reporting, Recapitulation, Hypotheticality, and Question. Failure to fulfil predictions results in text

appearing less coherent. Tadros' predictive labelling is characteristic of planned, deliberately structured, typically written, text more than negotiated, spontaneous, typically spoken, text. Similarly, Winter (1977; 1992) proposes the category of 'Vocabulary 3' for lexical items that organise text, and Francis (1986) describes the role of "anaphoric nouns". Flowerdew (2003) synthesises these proposals and notes that 'signalling' nouns ("any abstract noun, the meaning of which can only be made by reference to its context" (p.329)) are frequent, particularly in academic genres which are "removed from the concrete world of experience" (p.343), and are often cataphoric, or "forward-looking". Thus, on a more general level, text can be analysed as being frequently prospective.

2.7.5 Prospection and Encapsulation of Discourse

Sinclair (1994) claims that "each successive sentence is, for a moment, the whole text. This could lead eventually to a really operational definition of a sentence." (p.19). This view of text is explored in detail as Sinclair (1993) proposes the centrality of encapsulation and prospection to coherence in text (see Fig. 2.16). Encapsulation is a semantic process whereby the meaning of the previous sentence is included as part of the current sentence. It is this meaning that is carried forward in the discourse, rather than any component part of the sentence. Encapsulation occurs as each sentence in a text replaces the previous sentence as the current state of the text, including all of the developments in the text up to that point; each sentence can stand for the text so far and encapsulates the meaning of the preceding discourse by encapsulating the preceding sentence (which encapsulates the previous sentence and so on). Encapsulation is performed most frequently by 'logical acts', realised by signals of semantic relations (*so*, *therefore*, *and*, *yet*, *in fact* etc.), and by 'deictic acts', such as *this subject*, *this*, and *it*. While logical acts mainly refer to conjunctive relations, deictic acts closely match cohesive devices for participant tracking.

Prospection "leads the addressee to expect something specific in the next sentence" (Sinclair, 1993 p.12), although the expectation may be met after a delay, common in written texts, or not at all, as is common in casual conversation. Failure to fulfil the prospection results in a reduction in coherence. Prospection is most obvious in a question-answer sequence (and presumably in other adjacency pairs). A question prospects to the answer; the answer does not 'refer back' to the question, but fulfils the prospection that the question sets up, and thereby creates coherence. Similarly, the categories proposed by Tadros (1989) are also covered by prospection.

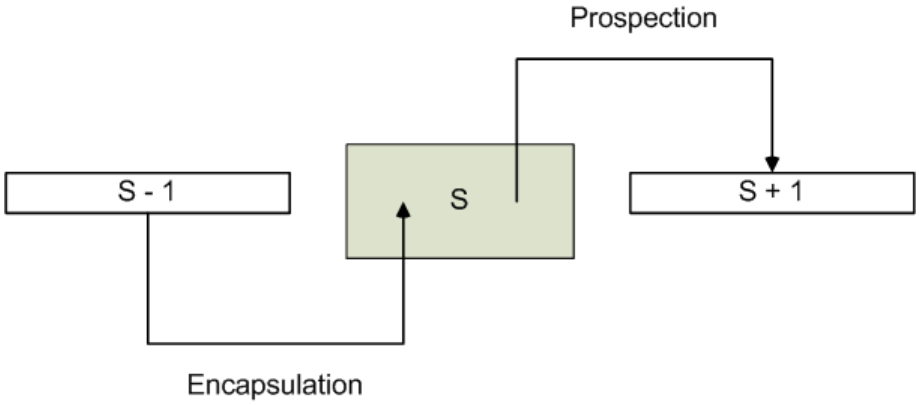


Fig. 2.16 Encapsulation and Prospection

Cohesive devices which work at the level of word or group are consequently considered less significant than the encapsulating and prospecting logical and deictic acts in creating coherence between sentences. Sinclair (1994) questions the importance of retrospective referencing, by depicting the processes involved in spoken language thus:

> People do not remember the spoken language exactly and so they cannot refer back to it in quite the simple way that they can in written language. Because we have the written text in front of us to check on, it is apparently easy to rely on retrospective reference. But do we really in the normal course of reading actually check back pronominal reference and so on? I doubt it. (p.16)

Although cohesive links, especially lexical ties and nets (Hoey, 1991b), are common, they are only significant in this model as part of the least common category of encapsulation, namely 'verbal echo'. Verbal echo is a repetition or replacement of a word or group, not a sentence as required by encapsulation. Lexical repetition and pronominal reference can act to create superficial cohesion, but not coherence. Sinclair concedes that a range of processes are likely to be involved during the processing of text, and so wishes mainly to underplay the role of cohesion rather than reject it.

The main difficulty with adopting Sinclair's model of encapsulation and prospection, however, is that, as a preliminary proposal, there appear to be no details, and we are offered no way of verifying an analysis. A more reliable system of analysis is essential to ensure reliable results. One way to achieve a more rigorous model of analysis is to focus on those elements that offer cohesive ties to elements in other sentences. The next section outlines one such model.

2.7.6 Prospecting Participants

The previous studies focused on prospection, and encapsulation, at the level of discourse. This section will now look at the validity of a Prospective approach to Participant Tracking – an apparent contradiction, but one which will be seriously examined, and could be easily resolved with a more appropriate label, such as Participant Continuity.

It is possible, following the proposed model of encapsulation, to view participant tracking with a greater emphasis on prospection than usually accepted. In order to track a participant it is necessary to make each participant unique – to distinguish each participant in the context or co-text. If we are to accept that each pronoun could refer to any possible participant in the text or context, as many formal approaches appear to, we have enormous potential for ambiguity. If, on the other hand, we accept Sinclair's (1994) proto-definition of a sentence as the current state of the text, we can find any possible referent by looking only at the current and preceding encapsulation. It is necessary only to identify the unique participants in each sentence in order to 'track' them – a far simpler process in a sentence-by-sentence analysis. If each sentence encapsulates the previous one, or prospects to the next, then the participants in the sentence will all be identifiable in adjacent co-text argues Sinclair:

> A word of reference should be interpreted exactly like a proper name or a noun phrase. The reader should find a value for it in the immediate state of the text, and not have to retrieve it from the previous text unless the text is problematic at that point. (1994 p.16)

Through encapsulation, the pronoun can be considered as being carried forward from a previous participant. In this case, the pronoun represents its previous meaning – it is an example of repetition. This suggests a very simple model of processing, but one that may not be able to function in some contexts, including literary texts. Within this model participants will only be tracked between two sentences, and will also only be a small part of the encapsulation and prospection of the whole sentence.

Emmott's (1994; 1997) model of prospection provides a more central role for participants. Using features of narrative to dispute Halliday and Hasan's (1976) view of locating an antecedent as a retrospective process, Emmott (1994) argues that readers of narrative set up 'frames of reference' which contain characters (or 'actors') in a specified space-time, and which remain constant until the writer explains otherwise. Emmott (1994) describes attributes of narrative text that cannot be explained fully by 'standard' reference theory: flashbacks – text that returns to a point after an interruption to pronominalise the

same actors as before the interruption rather than those mentioned last in the co-text; long gaps between mentions of actors; and covert participants – participants not named in text but understood to be 'present' in the context. These apparent anomalies point to a construing process required of readers quite different from a 'search back' through a text to find the previous mention of an actor, a database model of knowledge, or a dependence on the preceding encapsulation.

Detailing how prospection occurs in text, Emmott (1994) describes how actors, or participants, are 'carried forward':

> The frame monitors which characters are involved in the current action, bringing this information forward to each new sentence so that it can be used in the interpretation of such pro-forms … Indeed many readers seem to carry forward their mental constructs as a quasi-visual image (1994, p.159, emphasis added)

The 'frame' is a mental construct that allows the reader to envision the context described in the narrative. Emmott then explains that it is through the use of frames, rather than a search back through the co-text, that pro-forms can be interpreted. A frame contains a number of entities and participants. These participants are then 'primed' (Emmott, 1997) for the reader – the reader knows that they are part of the context, and is ready to focus on any one of them. When the writer focuses (Sidner, 1979) on one of the participants, that participant is separated from the others in some way but all other participants remain in the frame, and can all be potentially brought into focus, until they are explicitly removed from the frame (similar to the 'exit' of stage directions).

Emmott (1997) is more explicit about the role she sees for prospection, noting that centering theory and other Artificial Intelligence theorists are inconsistent when discussing forward- or backward-looking reference. Emmott reverses the typical approach to antecedents by describing participants as "anticipating a co-referential pronoun" (1997, p.221), and defines focusing as

> bringing the entity representation of a character forward as we read and remaining conscious of it as the focus of attention during processing of the immediate text. (1997, p.224)

Emmott (1997) critiques the standard retrospective view of anaphora by pointing out that trying to find the antecedent for a pronoun is like watching a film in reverse, as a reader will always come across the antecedent, or referent, first and the pronoun later. That is, the question should not be "Where is the antecedent for this pronoun?", but "When am I going to see this participant changed into a pronoun?" We could name each new

participant an 'anticipant' – a participant that is potentially a focus when it may be pronominalised. Emmott's approach differs from that of Halliday and Hasan (1976; 1985) by attempting to account for the psychological process in which the reader is engaged in order to describe the nature of anaphora and text. Halliday and Hasan take as their only data the text itself and what that reveals about anaphora.

Although Emmott's work so far has focussed on narrative, it seems likely that the same processes will be employed when reading non-narrative text. More likely than not, the reader will use whatever strategy is easiest, and so while there may be instances of tracing a pro-form back to its antecedent if it is not far away in the text, it is more likely that a process similar to that available to oral narrative can be employed to 'monitor' actors in all genres. That is, even without the evidence of the written text in front of us, we can still follow entities, participants and actors while listening to a text, and will still respond effectively to interruptions, flashbacks and so on. Since we are able to do this, and prospection and frames so far seem the most likely explanations for how we do it, there seems no reason to suggest that for non-narrative written text we would employ a more complex process. In fact, it is likely that frames are easier to construct when the reader is able to relate the entities and relations to schemata (Rumelhart, 1980) in their own experience, such as in expository text, rather than when the reader has to create, develop or adapt frames, as in most narrative. Ultimately Emmott (1997) is arguing for a model of analysis that emulates a model with maximum economy of processing effort on behalf of the reader/listener.

If we are to accept a prospective model, two points need to be made. First, prospection is being proposed here as a way of modelling the dynamics of the construction and interpretation of text, not as the only way of analysing textual phenomena. That is to say, Halliday and Hasan (1976) is a synoptic, in-vitro model for textual analysis, and does not intend to simulate the psychological processes involved. This study attempts to integrate the dynamic process into discourse analysis by considering research related to SFL. Second, that prospection does occur need not imply that retrospection does not. I would propose an analogy with Sinclair's (1991) notions of the 'Idiom Principle' and the 'Open-choice Principle'; prospection, like the idiom principle, is the easier, dominant, or unmarked, process while retrospection, like the open-choice paradigm, comes into effect when the text is 'problematic'. Retrospection is employed only when there are difficulties, potential communication breakdowns or new situations that make greater demands of the listener/reader. In effect, the model proposes that 'anticipants' are carried forward in

discourse until ambiguity or conflict is encountered in the text, when one strategy for resolution of meaning is the retrospective searching for participants.

2.8 Implications

In many descriptions of Reference (*e.g.* Bos, 2003; Gundel *et al.*, 2003; Lambrecht, 1994) one system is believed to be able to account for all of the variations. In the model presented here a distinction is made between the semantic relations of Reference and the location of phoricity. In the analysis of this corpus these two systems, of participant Identification and Tracking (Martin, 1992), reveal a wide range of variation in choices that contribute to the meaning-making resources of the textual metafunction. They appear to exhaust the choices available to the system of Reference, while helping to explain the apparent discrepancies in other descriptions. It would appear that more analyses that use the systems proposed by Martin need to be carried out. I predict analyses using this model will produce more accurate descriptions of reference in discourse.

The model in this study makes very little change to Martin's (1992) original formulation, but does propose the replacement of the category of Bridging with a hierarchy of semantic relations, originally proposed in Martin (1992) as part of an Ideational taxonomy, and based on Lyons's (1977) semantic relations (see discussion in 5.2.3 and 5.2.4). The application of this model to discourse analysis has revealed it to be a very powerful tool in identifying a range of semantic relations, including so-called 'direct' reference. The analysis has demonstrated that, for these texts, it can specify the different relationships identified as Bridging (e.g. Haviland and Clark, 1977), "closely connected" (Grosz, 1978), semantic distance (Almor, 1999), Inferable (Prince, 1981) and Mediated (Strube and Hahn, 1999) by using a network of semantic relations. The necessity for the concept of Bridging is contested, as Bridging asks more questions than it answers by assigning the processes involved to a psychological 'black box'.

Using a tool such as WordNet (Cognitive Science Laboratory, Princeton, 2005), it may be possible to partially automate the network of semantic relations. This would go some way to establishing whether a systemic-functional approach to the 'resolution' of anaphora (and related phenomena) would be at least as successful as formal-based models. It would certainly call into doubt the need for the category of 'bridging'.

Participant tracking does not allow the analyst to identify cohesion produced by non-nominal aspects of the clause, and does not analyse parallelism across clauses. While these aspects of the textual metafunction no doubt operate in text, they are presumed to be far less frequent and therefore have less effect on the structuring of information than the participant identification and tracking as analysed here. The model presented here is synoptic rather than dynamic, but it not does not preclude a prospective view of participant tracking because it is a model based on the construal of meaning within context. It is possible that when a new participant is Presented it becomes an 'anticipant', with the expectation that it will become pronominalised. A more dynamic model would benefit the theory in general, and so further research is required here.

Within the textual metafunction, participants play a significant role in creating cohesion and identifying the different parts of a message that may be carried forward in discourse. Participant Identification and Tracking account for many of the referential descriptions that other theories combine with information structure (see section 4.1). This chapter has attempted to demonstrate through text analysis that the systems of participant Identification and Tracking are distinct both from each other and from information structure.

2.9 Conclusion

The main aim of this chapter has been to establish a model of Participant analysis that is distinct from the other two systems of Theme and information structure, as proposed by Fries (2002). The discussion earlier in the chapter (section 2.2) attempted to define a participant in discourse analysis, and to search for related concepts in adjacent fields of linguistics. This search demonstrated the large gap in concept between the psycholinguistics-based theories, inspired by Clark, and the Systemic-Functional approach. Between the two approaches there is some common ground, however, to be found in the theory of Centering, which draws on computational analyses, and the work of Sinclair and Emmott, among others, all of which emphasise the importance of Prospection. While the importance of Prospection was identified, it has not yet been integrated into the analysis.

The definitions of participant in Martin (1992) and Halliday and Matthiessen (2004) allow for fairly quick classification. The analysis suggests very strongly that participant Identification and Tracking are two separate systems that perform quite separate

functions within the discourse semantics of the textual metafunction. The analysis also demonstrates that there is no need to posit a representationist viewpoint of reference; participants can be tracked in text regardless of their representational value. The sample texts demonstrate that through predictable semantic relations between referentially-tied participants, instantial relations and identities of participants are construed in text.

A number of issues currently remain unresolved, and merit further research. These include, as mentioned above, the integration of a prospective and, ultimately, dynamic perspective into the model. Further parallels with Centering studies should be explored, and the issue of multimodality must be examined in relation to the model in order for it to be applied in a pedagogical context. A study is also needed that addresses the hypothesis of the hierarchical view of semantic relations. This could be validated both textually, by allowing for more than one relation per participant, and psychologically by timing how quickly subjects can spot the 'odd one out', or its opposite, the item with a connection. Finally, while comparison between Systemic Functional Linguistics and schools of Psycholinguistics have been contrasted, a more in-depth discussion of the issue of representation and the construal of meaning in the two models is required in order to clarify the different positions in relation to reference theory. In conclusion, as far as the main aim of this study is concerned, the next step is to analyse the role of Theme in the textual metafunction, and to see how it interacts with the systems identified in the chapter.

Chapter 3 Theme

3.0 Introduction

This chapter investigates Theme – one of the three major systems within the Textual metafunction that operates within the clause. A discussion of the concept, delimitation, special constructions and the analysis of Theme both within the clause and within text, relying on research within Systemic Functional Linguistics (SFL) and 'neighbouring' schools, reveals a clearer picture of Theme, allowing for clearer comparison with the other systems in the Textual metafunction. Sample analyses are then provided to be included in a larger quantitative account of Theme. Finally, the quantitative data will provide a starting point for textual comparisons between thematic and participant options. This discussion will pave the way, theoretically and analytically, for the next chapter which will focus on the final clause-internal system within the textual metafunction – information structure.

3.1 What is Theme?

Of the three systems in the Textual Metafunction (Participant, Theme and Information), Theme has attracted the most research within SFL in a variety of genres and registers (*e.g.* newspaper: Francis, 1989; academic text: Gosden, 1993; North, 2005; Whittaker, 1995; radio and television talk: Bowcher, 2004; Gómez, 1994; Ravelli, 1995; and narrative: Cummings, 2005; Martin, 2000) and across different languages (*e.g.* Finnish: Mauranen, 1996; Chinese: Fang, McDonald and Cheng, 1995; French: Caffarel, 2000; German: Hasselgard, 2004; Portuguese: Vasconcellos, 1992; Spanish: McCabe, 2004; Taboada, 2004). Its analytical reliability has, therefore, been established by repeated study and through periodic reviews of the research (Butler, 2003; Fries, 1995a; Fries and Francis, 1992; Gómez-González, 2001b; Thompson, 2007). This study will not investigate Theme as broadly as the chapters on Participant Tracking and Information structure as a result of the wide range of studies and reviews that are already available, and because very few other theories have incorporated the concept of Theme into their linguistics models. What follows is a brief review of literature that will lead to guidelines for analysis of the data in this study.

Despite the wide range of studies, Theme is not uncontroversial. Debate has arisen around the meaning of Theme, its identification, particularly the limits of Theme in a

clause, special thematic structures, and Theme's role across clauses. These debates need to be reviewed before specifying the categories and criteria to be applied in the current research. We will start with the debate surrounding the meaning of Theme.

3.1.1 Theme is not just about what the clause is about

A great deal has been written on what may be considered unfortunate wording in Halliday's early formulations of Theme. Halliday (1967b) characterises the function of Theme as "what is being talked about, the point of departure for the clause as a message" (p.212). Criticism of Theme has often focused on the first part of this formulation (Huddleston, 1988), ignoring the importance of the second part (see Martin, 1992b; 1995a; and Vasconcellos, 1992 for responses). Downing (1991), for instance, notes that the Theme is frequently not what the clause is about, but it is always the starting point of the clause. The use of 'about' may provide a focal point for criticism, but it needs to be examined in the context of SFL theory in order to appreciate its implications.

In SFL, labels and terms are not arbitrary, but are motivated by their function. The meaning and function of Theme, as it is commonly understood, can be an underlying meaning that continues through a text. A theme of a clause, then, leaves traces of itself through the clause. It is a starting point – which we could call the *ground* of the clause – as well as being related to the development of the clause. This is not the same as saying that the clause, particularly the verbal group, is *about* the Theme. The use of *about* here prioritises one kind of meaning – experiential, or that reflecting our material reality. In some cases, the clause may exclusively be about its experiential content, which can be used as its only grounding. When a clause-initial element, the logical subject and a participant that has been tracked through discourse are realised by the same wording, and no other initial elements are included in the clause, then, and only then, can the Theme be said to reflect what the clause is *about* on an experiential level. Following Halliday's (1967b) predictions, various studies (e.g. Gómez-González, 2001a; Bowcher, 2004) have suggested that this is the unmarked case, but it is not the only case.

A Theme may not always be *about* the experiential matter of a clause, but in every clause the Theme is what the clause is *about* on a textual level; what the clause-as-a-message is *about*. A message can be *about*, or grounded in, the relationship between interlocutors, the context of the message or relations between parts of the message, or an object or entity in the context. Potentially, therefore, a message has, respectively, an Interpersonal, Textual

or Ideational Theme. These Themes can also be combined to produce a Multiple Theme (Halliday and Matthiessen, 2004), revealing again how an experiential Theme represents just one type of *aboutness*. If we only consider the notion of *about*, it is clear that the common understanding of the term, based in experiential meaning, is insufficient to describe Theme, but must be extended to interpersonal and textual meanings.

The emphasis on *aboutness* may be a result of other schools of linguistics imposing their view of language on Halliday's conception of Theme, rather than a fault with the original characterisation offered (Martin, 1992b). While the Experiential (or Topical) Theme is privileged in systemic analysis, in as much as it is taken to be the final, obligatory Theme (Fries, 1995b; Halliday and Matthiessen, 2004; Martin, 1992; McCabe, 2004), it is still only regarded as one of the possible Themes in the clause – as not exhausting the "thematic potential" of a clause. This is a major difference between systemic functional linguistics and other linguistic theories that prioritise ideational meaning, even to the exclusion of the other metafunctions as the result of a reliance on a predicate logic that demands truth conditions in order to generate syntax. This fundamental difference in perspective may explain why Theme has been misunderstood, even misrepresented, by other non-functional linguists. A clause by definition consists of experiential, or representational, elements, but the clause constituents that would most often be analysed in formal linguistics as Topic and Comment (e.g. Chafe, 1976; den Dikken, 2005; Göksel and Özsoy, 2003; Hajicova and Sgall, 1988; Jacobs, 2001; Sgall, 1975; 1987) only constitute part of the meanings that can be conveyed by Theme and Rheme when viewed as functional categories. That is, Topic within formal linguistics shares some similarities with Topical Theme (as well as sharing similarities with agent in an experiential analysis), but fails to identify Textual and Interpersonal Themes and so is unable to contribute to a (meta)-functional analysis of text. (In fact, Theme contributed to Halliday's (1964/1981) critique of the generativist notion that sentences with different surface features (e.g. sequence) but with identical grammatical constituents *mean* the same thing.) As Matthiessen (1992) explains, the Ideational metafunction demands the sequencing of elements, but it is only through the enabling role of the Textual metafunction that the sequence is given value and, therefore, meaning. For a full systemic functional analysis, therefore, it is necessary to analyse the contributions that all three metafunctions make to the meaning of clause and text, before the meaning of Theme can be fully understood.

Another criticism that has been made against Theme is the linking of starting point, or ground, with sentence-initial position. Halliday has frequently pointed out that in English,

and probably in many other languages, Theme is realised, but not defined, by sequence. In Japanese and Tagalog, for instance, affixes can be used to realise the Theme of a clause (Rose, 2001). Theme, "grounding", or "starting point of a clause-as-message" do not necessarily coincide with initial position.

3.1.2 Clarifying Theme

An additional guide that can be used to further clarify the concept of Theme is to study it not just at the rank of clause, but below the clause, at the rank of group, and above the clause in the clause complex. In order to explain the difference between the systems of Theme and information structure, Halliday compares speaker and listener orientation:

> The difference may be summed up in the observation that, in dialogue, 'given' means 'what you were talking about' while 'theme' means 'what I'm talking about'; and as is well known, the two do not necessarily coincide. (1976, p.180-181)

Theme, then, is the starting point for the speaker. It is the speaker's *here-&-now*. This observation can also be applied to nominal groups:

> [the] principle which puts the Theme first in the clause is the same as that which puts the Deictic first in the nominal group: start by locating the Thing in relation to the here-&-now – in the space-time context of the ongoing speech event. (Halliday and Matthiessen, 2004, p.323)

The same thematic ordering, from the restricted here-&-now to the unlimited development of the co-text, is also noted in the ordering of the verbal group (Halliday and Matthiessen, 2004, p.336). The meaningful contribution of Theme can also be seen in the sequencing of clauses in a clause complex. The role of an initial subordinate, or hypotactic, clause, (a "regressive sequence") uses thematic resources to "set up the local context in the discourse for the *a*-clause: they re-orient the development" (Halliday and Matthiessen, 2004, p.393) (see also Thompson, 1985). Theme, then, is the major principle for anchoring the message within a context, thus emphasising its significant contribution to the enabling role of the textual metafunction.

A number of metaphors, many of them spatial in nature, have been proposed to characterise Theme (Matthiessen, 1992). A common metaphor is that of a framework for the Rheme (Fries, 1995b; Gómez-González, 2001a). Downing (1991) develops the framework metaphor into a range of functions, as summarised in Table 3.1. Matthiessen (1995) takes the spatial / framework concept one step further by relating the textual function of Theme to its ideational function in text: "The Themes are selected in such a

way that they present the current logogenetic growth-point in the ideational system." (p.27) That is, although the textual metafunction can be analysed independently, its main function is to enable the Ideational and Interpersonal metafunctions, and so Theme functions as the current ideational growth point – a growth realised in the Rheme of the clause. The crossover between the textual and ideational metafunctions described here mirrors that described in the previous chapter where participant tracking, through a variety of cohesive relations, was used to explain the mechanism of bridging. We shall return to the metaphor of growth, or a dynamic development in the text, when we look at the function of Theme across clauses.

Function	**Metafunction**	**Realised by…**
1. to signal speaker's angle in relation to Mood	Interpersonal	choice of Mood
2. to set the main semantic (temporal, spatial, individual, situational) framework	Experiential	temporal, spatial, individual, situational elements
3. to indicate the logical direction in relation to preceding text	Textual	conjunctions, conjuncts
4. to indicate speaker's attitude		modal disjuncts
5a. to relate content of message to addressee or to context		vocative, relational disjunct
5b. to relate content of message to something else outside		relational disjunct
6. to signal change of speaker or change of direction in discourse		continuative

Table 3.1 Function of initial elements, based on Downing (1991 p.128-9)

As well as operating at different ranks (group, clause, clause complex and text), Theme can be analysed at different levels of delicacy to reveal greater detail in analysis; the research goals of the analyst will determine the degree of delicacy (Thompson, 2007). Within this study for instance, the t-unit proposed by Fries (1994; 1995), and used subsequently in other studies (e.g. Mauranen, 1996) to investigate the Method of Development, is not considered sufficiently delicate to identify the meaningful patterns created by the interaction of Participants, Theme and Information within a clause, as it identifies just one Theme for each clause complex. (It may also be suggested that since the T-Unit was intended as a measure of clause complexity for determining the development of student writers (Hunt, 1965), it may not be the most suitable unit for determining Theme and Method of Development.) One of the main research goals of this study is to examine the interaction of the systems of Theme, Participant and Information, and so a high degree of delicacy in the analysis of Theme will be required. Themes will be

identified and analysed for their functional contribution to the clause, but the Thematic role of each element within Theme will also need to be identified so that a comparison can be made with other systems. For example, a nominal group, including all of its dependent pre- and post-modifying elements, will be accorded the same status in an analysis of Theme. The internal thematic structure of groups will not be analysed in this study, as it seems unlikely that this degree of delicacy will impact on the results of the analysis in the present study. However, the component participants analysed separately as part of the participant identification and tracking analyses will later be compared to the analysis of Theme to examine the interaction of the systems.

Halliday and Matthiessen (2004, p.393) demonstrate a two-layered thematic analysis for a clause complex that reveals the thematic patterns both within single clauses and across the clause complex. In this study, Theme will be analysed prior to revealing the textual patterning of groups and clauses within text, particularly in relation to Participant Identification and Tracking. This means that the clause complex will not be analysed as having one Theme (*cf.* Fries, 1994), but the Theme for each clause will be classified. The need to analyse at this level of detail results in the proposal that all clause-like structures will be analysed for Theme and Rheme. Non-finite clauses, which carry no realisation of Mood, for instance, will be divided into Theme and Rheme (Halliday and Matthiessen, 2004, p.100). Assuming that Theme and Rheme are realised by sequence in English, it seems likely that sequence will have an effect on non-finite as well as finite clauses. Theme-Rheme analysis will be as exhaustive as possible so as to identify where all groups are placed in relation to Theme and Rheme.

3.2 Delimiting the Theme

Another discussion that is frequently referred to in research into Theme is the length of Theme in the clause, or how far the Theme can extend into the clause. Since the Textual metafunction organises meaning through 'waves', not through the constituent patterns associated with Ideational meanings (Halliday, 1979; 2002; Matthiessen, 1992), Theme is not defined in terms of syntax or constituency, and is not realised by syntactic constituents in the clause (see Fig. 3.1). This fact has contributed to the discussion surrounding how far one should look in order to find the Theme of a clause, bearing in mind that Theme can have multiple meanings (Textual, Interpersonal and Ideational) and can be realised by a combination of exponents of these metafunctions. We shall look

briefly at the outline of this debate before examining other aspects of Theme that will allow a working hypothesis for this study.

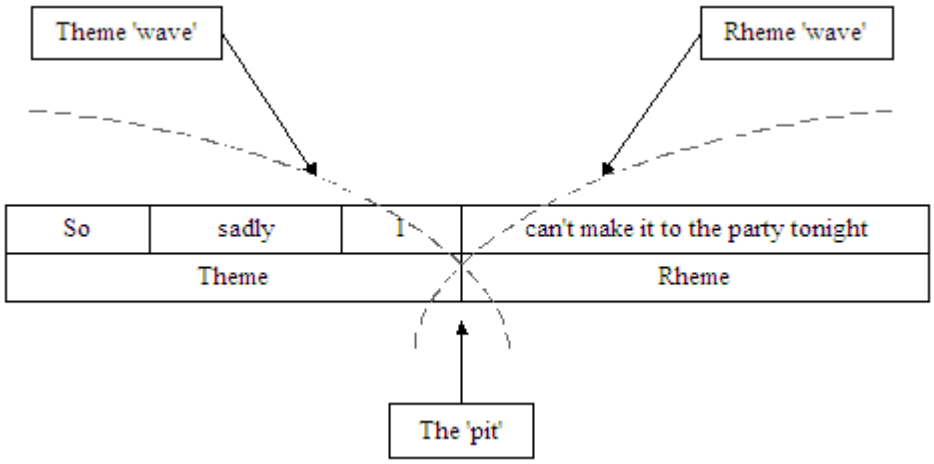


Fig. 3.1 The wave-based nature of thematic meaning in a clause

For some researchers, the Theme – the ground on which the clause is based – is restricted to the first element in the clause (*e.g.* Bloor and Bloor, 1992). One reason for this minimal approach could be that applying Theme and information structure to a pedagogical setting requires the simplest approach. Halliday's approach is to try to limit the extent of Theme. Theme for Halliday extends only as far as the first constituent in the experiential structure of the clause, whatever that constituent may be: "the Theme contains one, and only one, of these experiential elements" (Halliday and Matthiessen, 2004 p.79). Halliday's identifying criteria have also served a great many other researchers (Bowcher, 2004; Lassen, 2004; Leckie-Tarry, 1995; Martin, 1992b; McCabe, 2004).

At the other extreme, Berry (1995; 1996) tends towards a more extravagant analysis, including all elements in a declarative clause up to the verbal group. This disagreement stems from the assumption that the thematic potential is not exhausted until the (interpersonal) Subject has been included in the analysis of Theme. In a clause with an initial Circumstance, it is suggested that the thematic potential has not been exhausted, and that the subject still retains a degree of importance in the structure and development of Theme. Other researchers that consider the subject to be an obligatory part of the Theme include Caffarel (2000), Eiler (1986), Gosden (1993) Mauranen (1996) and Rose (2001), in some cases as a consequence of applying the concept of Theme across languages.

One approach to identifying the limit of Theme in a clause is to take a dynamic perspective, such that it is not possible, or even necessary, to identify the extent of Theme until the onset of the Rheme (Ravelli, 1995). Although it may be possible in a synoptic analysis to draw a line between Theme and Rheme, during real-time processing of text, whether written or spoken, a degree of ambiguity is perfectly acceptable, even to the end of the clause complex. There may, in fact, be good rhetorical reasons for intentionally suspending the identity of Theme. Rather than analysing for Theme and naming the remaining portion Rheme, Ravelli makes Theme and Rheme mutually defining by proposing that Theme can only be recognized when the onset of Rheme has been identified. Ravelli (1995) avoids a circular definition (i.e. that Rheme is what Theme is not) by suggesting that the "pivot point" of a Theme-Rheme analysis in English – the place where Theme turns into Rheme – is found at the finite/predicate, since it is only here that the Topical Theme (or subject) can be categorically distinguished from other potential experiential elements at the onset of the clause. A potential challenge to Ravelli's (1995) dynamic analysis, however, is that it only applies to declarative clauses in English. Further work needs to be done to define the 'pivot point' in other moods and languages. Since this corpus contains very few instances of Imperative or Interrogative mood, we will postpone this discussion. Thus, Ravelli (1995) concurs with Berry (1995), but offers a different rationale; the identification of Theme depends on the onset of the verbal group to locate the Subject of the clause.

One challenge that is raised by 'extravagant' approaches to Theme is how to deal with any remaining experiential elements within the Theme zone. That is, for Halliday only one Experiential element can function as Theme, because it exhausts Theme. If we are to include as part of our definition of Theme an obligatory subject, the status of any remaining pre-verb group experiential elements is open to question. In this study, which adopts the dynamic approach, Subject in Theme will be labelled Topical Theme, while any other experiential elements in pre-verb group position, apart from Subject, will be labelled Experiential Theme. This allows for more than one experiential element to be considered part of Theme. As this does not match Halliday's definition of Multiple Theme (one Experiential plus preceding Textual and/or Interpersonal elements), which allows for just one experiential element, this analysis will use the term Complex Theme in opposition to Simple Theme, which consists of a Topical Theme only. A Complex Theme may contain Texutal and Interpersonal thematic elements, following Halliday's terminology, depending on their metafunctional contribution to Theme

Extending Pike's (1959, cited in Matthiessen 1988) insight, Halliday (1979; 2002) expounds the view that the three metafunctions are realised by inherently distinct patterns of meaning. Using the metaphor of field, particle and wave, for Interpersonal, Ideational and Textual meaning, respectively, Halliday demonstrates how meaning in the three metafunctions is distributed through the clause. Martin (1992) and Matthiessen (1988; 1992) also discuss the implications of the alternating patterns of prominence and non-prominence in Textual meaning across clauses. Matthiessen (1992) applies this perspective to the debate of the extent of Theme by suggesting that it is not possible to draw a dividing line between the Theme and Rheme of a clause, as a result of the non-constituent nature of the textual metafunction. That is, since the textual metafunction operates in wave patterns in the clause and text, its boundaries are inherently indeterminate. Thus we find ourselves in a position where, instead of dissecting a clause into its Theme and Rheme, we can characterise a particular constituent in a clause as more or less Theme- or Rheme-like. Cummings (2005) describes the importance of the wave-like pattern to Theme thus:

> Failure to recognize the thematicity of a preverbal subject when it is preceded by another experiential element is to impose a categorical and segmental model of Theme somewhat too zealously on a graduated phenomenon (p.129)

Perhaps the safest option is to identify Theme 'zone' and the crest of the Rheme, while accepting that between the two there is an indeterminate 'grey' area. That is, there is in most clauses an area which is neither Theme nor Rheme, and has been termed the 'pit' (Hartnett, 1995). Therefore, in this study unless the part of the verbal group is in final position in the clause, it will remain unanalysed for Theme and Rheme – there being no compelling argument to force it into the characterisation of Theme or Rheme since it is strongly associated with neither.

Although there has been much debate on the extent of Theme in a clause, it is generally agreed that Theme is multifunctional with an unmarked sequence of: (Textual) ^ (Interpersonal) ^ Ideational[2] (Bowcher, 2004; Gómez-González, 2001b; Halliday, 1967b; Halliday and Matthiessen, 2004; Vasconcellos, 1992,). When there is only a Topical Theme, it exhausts all of the thematic potential. If the obligatory Topical Theme is preceded by other optional thematic elements, thematic potential is shared. The fact that there is some choice for placing Textual and Interpersonal meanings into Theme position

[2] This coding is commonly applied in Genre studies, where parentheses show optional items and a caret shows sequence.

suggests greater motivation, and possibly greater thematic meaning. Since a conjunction used as a Textual Theme has an obligatory first position in the clause, it has been suggested that it has less thematic meaning (Fawcett, 2003). Halliday (1994; Halliday and Matthiessen, 2004) points out that conjunctions are thematic, but they do not exhaust the thematic potential. An alternative, discourse rather than clause, perspective would suggest that the use of a conjunction is motivated by the choice to combine clauses into a clause complex rather than leaving them as separate clauses. This is clearly a meaningful textual choice.

Although the extent of Theme may vary in different studies, all of the studies have demonstrated some degree of validity through analysis of text. Each definition of the extent of Theme produces slightly different results in textual analysis (Thompson, 2007). Thus, taking a research-based view, the analyst should consider the aims of research before deciding on the limits of Theme. This study attempts to investigate the textual metafunction from a discourse stratum, taking a great deal of its theoretical support and analytical models from Martin's (1992) "English Text", and aiming for a dynamic perspective. In this chapter, Theme is analysed first independently, and then in comparison with an analysis of participant tracking. In the following chapter, these analyses are combined with information structure. As both Matthiessen's view of a wave-like pattern and Ravelli's dynamic approach take a text-based, rather than a clause-based, approach to Theme, they seem the most appropriate models of Theme for this study. In short, the study will adopt the more extravagant view; that Theme extends as far as the start of the 'pit' (Hartnett, 1995) of the verb group. The verbal group will then, unless it is clause-final, be considered as indeterminate in terms of Theme-Rheme structure. It represents, therefore, the corresponding dip of the Theme wave and the initiation of the Rheme wave.

3.3 Special Thematic Structures

A range of structures have developed in English, especially in the written mode, to provide greater flexibility in choosing thematic elements. These structures typically use an Identifying clause. In Identifying clauses the two parts of the clause, labelled Identifier and Identified, are ideationally equated, releasing them from the bounds of sequence, and allowing Theme and information structure to be arranged as per textual goals. That is, notwithstanding the resource of passive voice, it often happens that transitivity choices

force an element into Rheme position that would otherwise, for discoursal purposes, be chosen as Theme. Special thematic structures (underlined in the examples below) free discoursal choices from the confines of transitivity choices. Typical examples include:

3.i What a thematic equative does is to create a sense of exclusivity.

3.ii It was his teacher who persuaded him to come.

3.iii It was the waiter who told us to sit at this table that made the mistake.

3.iv It is unlikely that he will be here on time.

These structures, known variously as 'enhanced Theme' (Huang, 1996), or as cleft, pseudo-cleft (e.g. Prince, 1978; Collins, 1991) or it-extraposition (Kaltenbock, 2005) in formal grammars, use a combination of nominalization and rankshifting to reposition elements in the clause.

The first of these structures, known as a pseudo-cleft sentence in formal grammar, is a thematic equative (example 3.i). Halliday points out that Relational-Identifying clauses that include an embedded or rank-shifted clause (also known as noun clauses in traditional grammar) in the first part of the clause are frequently used to allow more complex ideas to be placed thematically. Thematic equatives

> have evolved, in English, as a thematic resource, enabling the message to be structured in whatever way the speaker or writer wants. (Halliday and Matthiessen, 2004, p.71)

In this structure, the functional role of the embedded constituent in an experiential analysis remains that of a participant (identifier or, more commonly, identified). The rank-shifted clause contains its own thematic structure, which may be analysed, but the analysis is different to the matrix clause in terms of delicacy. The rank-shifted clause also functions to create a sense of exclusivity (this one and no other) for one of the participants. This is true whether the rank-shifted clause is placed in Theme or Rheme.

A predicated Theme (examples 3.ii and 3.iii) shares, with thematic equatives, both the flexible structure of an identifying clause and the function of assigning exclusivity to a participant. It differs, however, in that it does not use a complete embedded clause to act as participant. In a two-tier analysis of thematic structure (see fig 3.2 below) Halliday presents an analysis similar to that for sub- or co-ordinated clauses (Halliday, 1994, p.60; Halliday and Matthiessen, 2004, p.97). That is, the element after the 'relative clause' marker (*who*, *that* or *which*) is analysed as a constituent of the clause complex.

Consequently, the hypotactic or embedded clause is analysed as having its own internal thematic structure of the same order as the analysis of the independent identifying clause (see layer (a) in Fig. 3.2). Another 'layer' of analysis classifies the predicated independent clause as Theme with the hypotactic clause as Rheme (layer (b) in Fig. 3.2). Example iv, although it may resemble predicated Themes, does not contain a relative 'that' clause, and so is not a considered a 'cleft' structure. However, it does follow the same pattern as being proposed here, with the analysis showing two clauses and two Themes, one being fulfilled by 'It' and the other by 'he' (see Thompson, 2007 for discussion).

	It	was his teacher	who	persuaded him to come
(a)	Theme	Rheme	Theme	Rheme
(b)	Theme		Rheme	

Fig. 3.2 Thematic structure of clause with predicated Theme (Fig. 3-22 in Halliday and Matthiessen, 2004, p.97)

One alternative to Halliday's analysis has been proposed by Huang and Fawcett (Huang 1996; Huang and Fawcett 1996; Fawcett, 2000) who argue that the "It + *be*" section of this structure should not be analysed as having an independent thematic structure, but that this constitutes the "thematic build-up" to the main clause. In the example in Fig 3.3, Huang (1996) only recognises "his teacher" as Theme. This is because, in Fawcett's model *it* is not referential and is therefore experientially 'empty', the verb *be* functions to enhance (or highlight) the Theme that comes after the verb, and because Fawcett recognises different participant roles (Carrier and Attribute) in Identifying clauses to those used by Halliday. This analysis is partly supported by investigating equivalent structures in Mandarin Chinese (Huang and Fawcett, 1996).

	It was	his teacher	who persuaded him to come
(a)	Thematic build-up	Theme	Rheme

Fig. 3.3 Huang and Fawcett's Thematic structure of clause with predicated Theme

Huang (1996) offers a taxonomy of discourse functions for enhanced Themes (Fig. 3.4), based partly on Prince's (1978) categories and resembling Prince's (1981) taxonomy of 'given-new' (see 2.2.3). In all cases, however, Huang's taxonomy (Fig. 3.4) appears to conflate the functions of Theme, Reference and information structure, probably as a

result of using Prince's cognitivist definitions of Given and New. It seems that many of the functions itemised in Huang's (1996) taxonomy, such as re-activating, initiation, correction, refer to Participant Tracking and information structure, realised by the tonic foot when read aloud, more than they do the functions of Theme. Furthermore, there appears to be little or no reason to object to an experientially empty "it" as Theme, when there are good reasons to analyse the equally experientially empty "there" in existential clauses as Theme (Martin, 1992b; 1995a; Mauranen, 1999). That is, the starting point of the clause for the speaker, may in fact be considered experientially empty or "structurally cataphoric", i.e. referring to an imminent structure as a whole rather than a particular participant (Martin, 1992). The analysis offered by Halliday and Matthiessen above (Fig. 3.2), also identifies *it* as the Theme of a predicated clause. This analysis lends further support to separation of Theme from other functions in the clause. If the predicated *it* is analysed as both interpersonal Subject and Theme, but is experientially empty, it is clear that Theme is distinct from the formal category of Topic. That is, the predicated *it*-Theme (along with existential *there*) is marked by creating a clause with no Topical Theme.

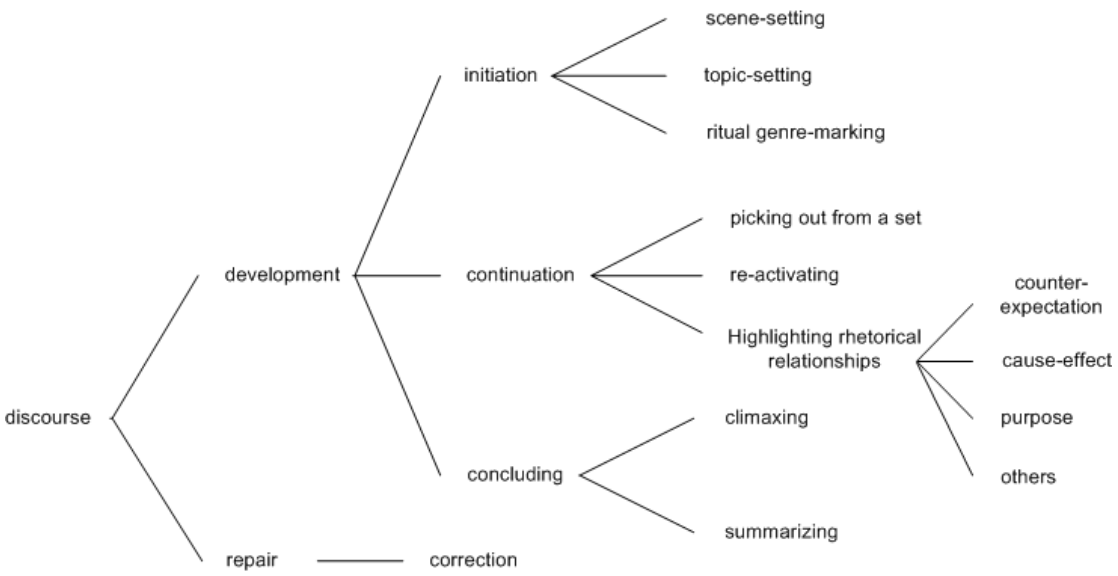


Fig. 3.4 Discourse functions of enhanced Theme (from Huang, 1996 p.93)

Although there are similarities between thematic equatives and predicated Themes, it appears that it is mainly information structure that is affected by predicated Themes (as described in Halliday and Matthiessen, 2004, p.95) and so will be examined in more detail in the following chapter (section 4.7). Halliday and Matthiessen (2004, p.643) also point out that these structures are often a result of grammatical metaphor, and the main beneficiary of the non-congruent configuration of meanings is the textual metafunction.

Any examples of predicated Themes found in this study will be identified for each ranking clause, following the (a) level of analysis in Fig 3.2.

3.4 Theme across Clauses

Having dealt with the issues of the meaning and extent of Theme in a clause, we can now focus on the issue that relates to the function of Theme across clauses. Although Theme is realised within the clause, it contributes to the textual metafunction mainly because of its role across clauses. In fact, it has been argued that Theme can only be understood in relation to its discoursal function:

> … the semantic description of Theme as the "point of departure" can be understood only in the context of textual organisation, for it is this aspect of language use to which patterns of thematic selection bear some non-random relation… (Hasan and Fries, 1995 p.xix)

This section will investigate and distinguish Thematic Progression, Method of Development and hyper- and macro-Theme, retaining those analytical concepts that will enable a clearer picture of the textual metafunction and information structure.

3.4.1 Thematic Progression

Thematic Progression assumes that discourse is connected by constituents that are repeated in predictable places within the clause. The three main patterns of Thematic Progression identified by Daneš (1974), in the theoretical framework of Functional Sentence Perspective, are Linear, Constant and Derived (see fig 3.5). Linear describes a pattern whereby an element from the Rheme of the preceding sentence is placed in Theme position. In a Constant pattern of Thematic Progression, the same referent is kept for a series of Themes. Derived Thematic Progression uses thematic and rhematic elements from a single initial sentence as Theme in subsequent sentences.

A number of studies have confirmed that these (proto-)typical patterns of Thematic Development demonstrate some validity (e.g. Daneš, 1989; Green *et al*, 2000). However, some analysts have pointed out that the patterns are not as neat as Fig. 3.5 suggests. Consequently, Dubois (1987) revised the categories, pointing out that frequently the patterns are carried out across non-contiguous sentences. Subsequent studies have confirmed the contrast suggested by Dubois (1987) between the 'contiguous' and 'interrupted' thematic patterns, using alternative terms: 'continuous' or 'floor-taking' in

dialogue (Romero-Trillo, 1994); 'continuous' or 'breaking' in translation (Hasselgard, 2004); and 'constant' or 'split' in sports commentary (McCabe, 2004), and 'continuous' and 'switching' based on studies of various languages (Rose, 2001), respectively.

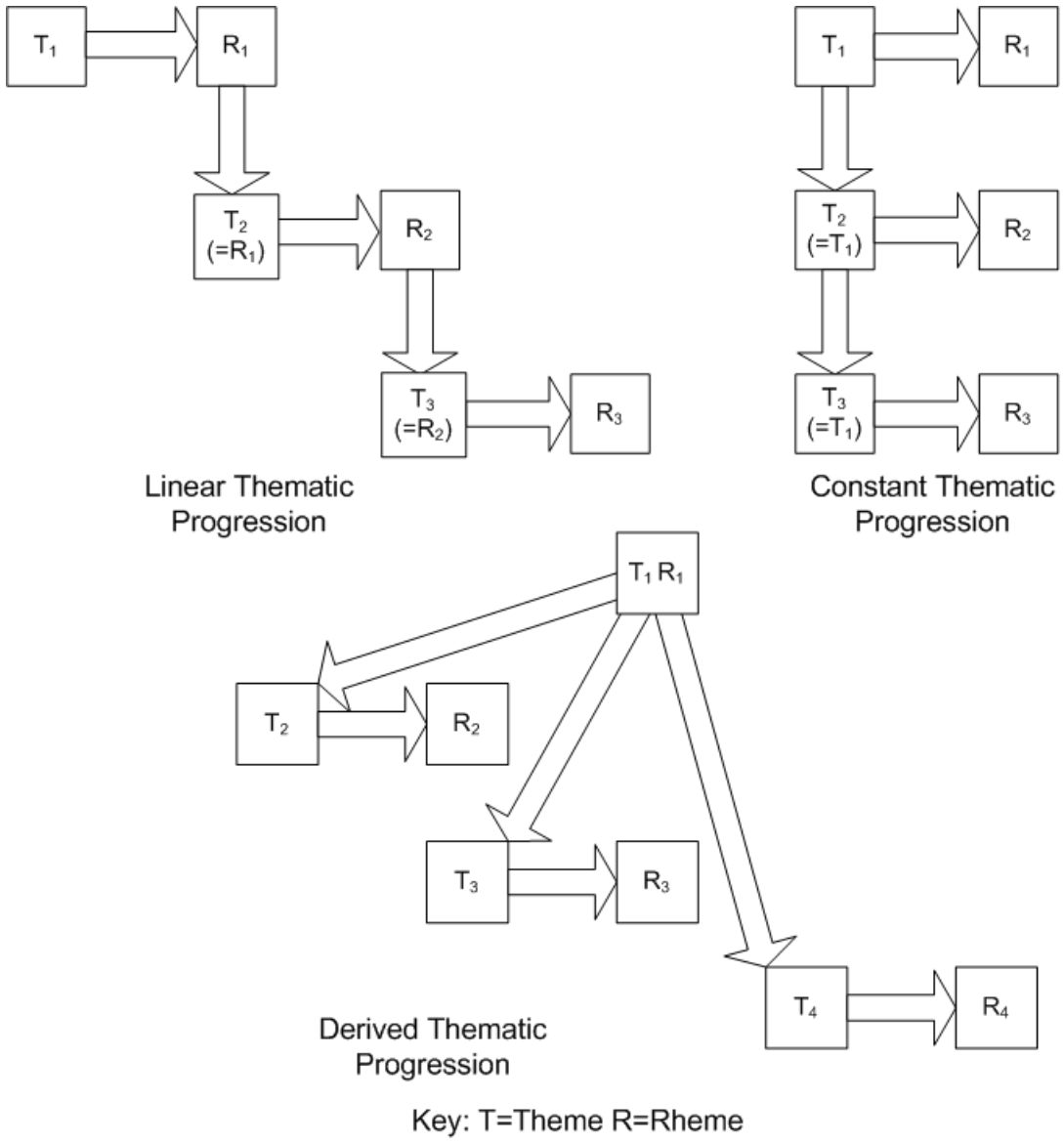


Fig. 3.5 Patterns of Thematic Progression (from Daneš, 1974)

The definitions of Theme in Systemic Functional Linguistics (SFL) and Functional Sentence Perspective (FSP) have major differences, however (Davidse, 1987; Fries, 1981; Halliday, 1974; 2002). One fundamental difference is that in FSP, as described by Daneš (1974) and Firbas (1992), the concept of Theme is not solely defined by sequence, but is also realised by cohesion, especially reference. Firbas' *Communicative Dynamism* (CD) employs a range of features to assign communicative status to an element in a sentence. Among those features are sequence and given, which relates directly to close-range co-textual reference. In FSP, Theme is the result of the combination of a variety of factors

that produce the lowest CD. All of the factors in communicative dynamism combine to assign thematic or rhematic status. That is to say, if an element is at the end of a clause but is the only Given (*i.e.* Presumed) information, FSP can categorise it as Theme because it has the lowest communicative dynamism. Clearly the same element would not qualify as Theme in SFL. Because the Theme and Rheme in FSP can enter into patterns of thematic progression, it must be remembered that the criteria for identifying Theme and Rheme in FSP and SFL are different.

From an SFL perspective, then, the main claim that can be made by Daneš and FSP is that participants are presented and presumed in consistent patterns across sentences. Although this may appear to trivialise a complex argument, it is perhaps this fairly broad interpretation that explains how such a wide range of registers and genres fit this narrow range of patterns. What is more surprising is that similar patterns have been found using an SFL definition of Theme (e.g. Fries, 1981, McCabe, 2004). What is needed, however, is a model of Thematic Progression that takes into account the SFL concept of Theme. This is precisely the gap that can be filled by Fries' Method of Development.

3.4.2 Method of Development

Fries (1981) demonstrated with a number of short texts that Theme, as described in SFL, can enter Daneš' (1974) typical patterns of Thematic Progression. Fries also introduced the term 'Method of Development', which he distinguished from Thematic Progression, to demonstrate how Theme can influence readers' perceptions of the whole text. In characterising Method of Development Fries appears to focus on ideas, readers and extended text:

> The way in which a text develops its ideas can be called the method of development of the text. The method of development of a text affects the reactions of its listeners and readers. There is no requirement that a good text develops its ideas in a single self-consistent manner. (1995a, p.9);
> [a single nominal topic, obvious point or Method of development] may be present to a greater or lesser degree. The strength with which the simple method of development and the single point will be perceived will relate to the contents of the Theme and Rheme of the component clause complexes of the text segment. (1995b, p.54-55)

For Fries, then, the Method of Development is very strongly tied to the perceptions of readers. As with Thematic Progression, it does not need to be present, and may be so to a greater or lesser degree. Consequently, it is not a formal requirement of a text; a text can function without a method of development. It is, perhaps, best seen as a stylistic marker.

A simple and/or single method of development will be perceived by readers as such – simple or following a single line of argument.

A number of writers on Theme have uncritically applied Method of Development, including Halliday (Halliday and Matthiessen, 2004), Matthiessen (1995), and Cummings (2005). In his review of research into Theme, Fries (1995) concludes that there has been limited support for an SFL-based view of Theme following Daneš' patterns of Thematic Progression, but that confirmation of Method of Development requires research that elicits reader responses to various texts that have been analysed with having an identifiable Method of Development. According to Fries, then, there is a clear distinction between the empirically-verified text-based Thematic Progression and the non-verified reader-based Method of Development.

There is some debate, however, surrounding the identification of the method of development, since Fries offers no clear guidelines for analysis. To look for a better understanding of method of development it may help to review how other researchers have developed this concept. Rose (2001), for instance, offers a semantic-based view, similar to Fries, when describing method of development as "the role of clause Themes in organising the method in which a text unfolds" (p.116). Cummings (2005) emphasises the importance of consistency in the method of development of extended texts, as opposed to the more local patterns between clauses.

The analyses offered by Fries (1981, 1995) to exemplify the method of development appear to prioritise ideational Themes. Caffarel (2000) develops this priority by claiming that, at least for French, the initial experiential elements in multiple Themes contribute to "the method of development of the text while the other ideational Themes have a more local function" (p.259). Mauranen (1996) also identifies the emphasis on experiential Themes in Fries' work:

> The method of development that a writer has used in a text passage, as suggested by Fries, can be discerned by looking at all sentence Themes and observing the continuity in their contents. By sentence themes Fries means the ideational constituents of the sentences, that is, the leftmost constituents which convey experiential meaning. (p.201)

Matthiessen (1995) prioritises thematic experiential elements in the method of development, when examining the enabling role of Theme in the textual metafunction. He describes the importance of Theme in organising the local relations between clauses

which exhibit the logical relations of Enhancement, Extension and Elaboration. These relations build the taxonomy under discussion in the text:

> The selection of Theme in a monologic passage is typically made to reveal the point of expansion. Thus if the expansion is one of temporal enhancement, the Theme is likely to be a specification of time; but if the expansion is one of taxonomic elaboration, the Theme is likely to be the current term in the taxonomy. (p.26-27)

Matthiessen uses Conjunction and method of development to explain the need for marked Themes. When the next Theme in sequence is not enhancing, extending, or elaborating the taxonomy of the text, a marked Theme is required to indicate the marked relationship.

So far in this discussion, the identification of method of development has been largely based upon the experiential content of a series of Themes in a text. However, some theorists have argued that this is an inadequate characterisation of method of development, if it is to provide any insights into the way that texts are organised across registers and genres. Although Fries (2009) objects to a definition of method of development, especially one which equates it with the collation of a sequence of experiential Themes, he has yet to propose a more thorough description of method of development, as demanded by Crompton's (2004) critique. Both Mauranen and Matthiessen stress the inadequacies of focusing on experiential elements in describing the method of development. Mauranen (1996) challenges the prioritisation of experiential elements in the definition of both Theme and method of development, as some languages, including Finnish, do not require the grammatical subject to take a pre-verbal position, eradicating the demand for an experiential element before the verbal group.

Matthiessen (1995) recognises that although taxonomies are the typical method of development of monologic text, necessitating a method of development dependent on experiential Themes, interpersonal Themes are frequently used in dialogue and may be used as the method of development. Pitjantjatjara is one example of a language which has not developed into a written mode (although it has been transcribed), and it uses a predictably higher proportion of Interpersonal Themes characteristic of the moment-to-moment negotiation typical of spontaneous conversation (Rose, 2001). Berry (1996) uses informant writers to align communicative intent with thematic content and concludes that where textual and interpersonal meanings are prioritised they are more likely to be realised in Theme.

Certain texts are more likely to select textual Themes to establish a consistent method of development, such as those typically described as 'procedural' as a result of a dependence on sequencing for method of development. Martin (1995a) makes it clear that Interpersonal, Textual and Ideational Themes not only contribute to a text's method of development but that some variations of method of development are valued by certain sub-cultures more than others (Martin, 2000). For instance, in high-school examination scripts, a method of development that does not focus on personal response in a review text receives a higher score. We see, then, a pattern evolving in method of development related to genre that was predicted by Fries (1981) but remains empirically unverified.

As a genre is realised in a particular register, the configuration of Field, Tenor and Mode vary the realisation of the method of development. Matthiessen (1995) analyses texts whose main function is to detail relations between inanimate parts to a distant audience. These tend to rely on Themes that describe the taxonomic relations, often depending on the parts themselves to provide the method of development. Martin (1992) provides an example of a text which attempts to "engage" its reader by employing an interpersonal method of development. Spoken conversation, where one of the functions of language that is typically highlighted is the relationship between interlocutors, consists of a larger proportion of interpersonal Themes which are used to advance the text. It is often difficult to identify a single method of development in conversation because Thematic choice is negotiated (Thompson, 2007).

Taking a discourse perspective, Martin (1992) characterises method of development by relating it to textual coherence and texture. The following quotation should suffice to summarise the discussion to this point:

> Method of development … establishes an angle on the field. This angle will be sensitive to a text's generic structure where this is realised in stages. Method of development is the lens through which a field is constructed; of all the experiential meanings available in a given field, it will pick on just a few, and weave them through Theme time and again to ground the text (p.489)

That is, Martin sees method of development functioning in the text in very much the same manner as Theme in a clause. He emphasises the importance of the experiential metafunction to generic structure, but as we will see in the next section, Martin also recognises the functions of interpersonal and textual Themes.

3.4.3 Hyper- and Macro-Theme

Martin (1992) emphasises the importance of method of development in establishing meaningful textual patterns. Although he recognises Daneš' (1974) patterns of Thematic Progression, he notes that the main function of derived thematic progression is to predict the method of development of (a section of) a text. He labels this hyper-theme, after Daneš, and proposes a further predicting Theme, called macro-Theme. The relationship is set up as follows:

> macro-Theme: text::
> hyper-Theme: paragraph::
> Theme: clause (Martin, 1992, p.437).

Martin then demonstrates how hyper-themes predict method of development with examples from written and spoken texts, although there is no attempt to identify the boundaries of a 'spoken paragraph' (see Pickering, 2004 for an attempt to do so). In reworked written texts, "the lowest-level hyper-Themes in a text are referred to traditionally as Topic sentences and the highest level "macro"-Themes as Introductions" (p.444). In prepared speech, many of the same features can be expected, with the speaker 'scaffolding' the structure of the text for the benefit of the audience. In spontaneous conversation, in contrast, the amount of planning is minimal. Hyper-Themes in conversation tend to frame shorter activity sequences, acting to "punctuate" the discourse rather than predicting it,

> …annotating the text in episodes as it unfolds rather than scaffolding it as a macro-constituent that is in some sense preconceived. Seen in this light hyper-Themes can be interpreted as a further category of textual metaphor: a hyper-Theme is metaphorical marked Theme (Martin, 1992, p.447)

That is, in spontaneous conversation, hyper-Themes frame a string of clauses in the same way that first-position subordinate clauses are said to provide the frame for the main clause (see "regressive sequence" in 3.1.2 above). Although speakers may not be able to predict the outcome of a conversation, at least partly because it is the result of interaction, they are in a position to suggest a framework for the ensuing monologic turn or dialogue, which may then be further negotiated and framed again by another hyper-Theme. Although this is an interesting proposal, I am unaware of any research to support hyper-Themes in conversation in either Systemic-Functional or Conversation Analysis literature.

Martin (1995b) demonstrates how the three patterns of particle, field and wave used as organising metaphors for ideational, interpersonal and textual meanings, respectively (Halliday, 1979; 2002; Matthiessen, 1988; 1992; see 3.2 above) can be extended from the

clause to the text. This application of Halliday's proposal to discourse uses the metaphor of fractal patterns of organisation; following a mathematical model, fairly simple fractal formulae frequently create complex, intricate patterns that are recognisably identical at different levels of detail, or delicacy: "principles of construing our experience of the world that generate identical patterns of semantic organization which are of variable magnitude and which occur in variable semantic environments" (Halliday and Matthiessen, 1999, p.223). Martin (1995b) describes how hyper-Theme and hyper-New in a non-academic text can then be reworked into a text creating a "hierarchy of periodic structure. The complexity of this hierarchy is in principle dependent simply on the amount of semiotic reality being organized" (p.29). That is, using a fractal metaphor, Themes, hyper-Themes and macro-Themes can be predicted or consolidated in increasingly-complex patterns, with meanings embedded inside other meanings through a text. Similarly, New, hyper-New and macro-New accumulate prior meanings in texts to further structure a text (Martin, 1992; see section 4.2.1).

A typical analysis of macro- and hyper-Theme is offered by Coffin (1997) who demonstrates how Themes are used in school history to scaffold arguments on a textual level in an essay. She contends that socialising students into the genre of assigning causality to historical events enables them to progress in their apprenticeship of academic history, and that it is mainly through the development of appropriate linguistic resources that the student succeeds in this apprenticeship. Macro- and hyper-Theme offer a powerful resource to scaffold arguments of causality:

> Located in hyper-Theme and combined with internal cohesive devices causality becomes a staging and ordering device. … This creates a cohesive text with each main consequence presented in the hyper-Theme both relating anaphorically to the macro-Theme and cataphorically to the 'mini' account sequence used in its elaboration (Coffin, 1997, p.218)

Coffin exemplifies this view with a school history essay that contains a macro-Theme which is elaborated by a series of hyper-Themes, which themselves are further elaborated. Ravelli's (2004) analysis also describes the role of hyper-Themes in the writing of management and history students.

3.5 Rheme

This short section is intended only to raise a few issues in relation to the present study, and to provide some areas for textual analysis that may reveal interesting results. Following convention in SFL, the Rheme will be left unanalysed in this study. In the system of Theme, Rheme is considered a 'residue'. SFL also uses a residue approach in other systems. For instance, SFL Mood analysis divides a clause into two parts. The first part, including the Subject and Predicate, indicates the Mood of the clause (*e.g.* Indicative, Declarative). The remainder of the clause is no longer relevant to a Mood analysis, and is labelled Residue. While there are certainly meanings being made in the Residue, they are not relevant to the analysis of Mood. Similarly, the Rheme may contain meanings central to Transitivity or to Information, but not relevant to an analysis of Theme.

As was noted in section 3.2, labelling Rheme a residue produces a circular definition – Rheme is purely what Theme is not. The main problem with the 'residue' approach is that under such a scheme Rhemes cannot exist on their own – if Themes define Rhemes, Rhemes cannot exist without Themes. Section 3.8 provides evidence that this may not be the case. If so, future research may be able to establish a revised definition of Rheme. While the 'residue' definition alone may be unsatisfactory, the fact that no theorist in SFL appears to assign a function to Rheme is more cause for concern. Halliday and Matthiessen (2004) for instance explain in detail how to identify Theme, but leave Rheme unanalysed.

Matthiessen (1995) describes Rheme as the location for the logogenetic growth of Themes. While this is a step forward, it does little more than explain that whatever appears in Theme will be developed in some way in Rheme – again a circular argument considering the function of Theme as the starting point of the clause-as-message: if Theme is the starting point, it must have somewhere to go, *i.e.* the Rheme. As Matthiessen (1995) and Martin (1992) note, the range of meanings in Rheme are far less predictable and far more varied than those in Theme. It could be conjectured that an analysis of Rhemes that included at least as much delicacy as a thematic analysis, including position and metafunctional role, would reveal predictable patterns. However, as there has so far been no other proposal to characterise the nature and functions of Rheme, this is not attempted in the present study.

Taking into account the earlier discussion regarding the wave-like nature of the textual meanings, it could be assumed that this is also true for Rheme, i.e. that the Rheme wave ends in a crescendo where the items are more Rheme-like than previous elements. However, as we shall see in the following chapter, this appears to be the same realisation for New information in written English, paralleling Fries' (1992) N-Rheme and Matthiessen's (1995a) Culmination. Similarly, it could be argued that Hyper- and Macro-Themes are reflected in Rhematic structure. In this study, the proposal that Information structure, rather than the Rheme, takes on the role of structuring longer stretches of text will be investigated in the following chapter.

3.6 The Logical Metafunction

Although this study investigates the textual metafunction, the logical, interpersonal and experiential metafunctions exert their influence at different times. This point in the analysis seems an appropriate time to analyse the logical structure of, and relationships between, the different clauses.

A major aspect of the logical metafunction is its ability to exert its influence over a wide range of ranks and under different conditions (Taylor Torsello, 1996). This is particularly true when discussing thematic choices in text because of the categories of Progressive and Regressive sequences (Halliday and Matthiessen, 2004, p.393; Thompson, 1985; see 3.1.2). A Progressive sequence is the relationship between two hypotactic clauses whereby the initial clause is the 'alpha' or independent clause and the second clause is the 'beta' or dependent clause. In a Regressive sequence the order is reversed. Furthermore, as one role of the textual metafunction is to assign prominence of various kinds, it is hypothesised that logical relations, such as hypotactic relations which also realise a type of prominence, will interact with this system. For instance, it is hypothesised that if all of the textual systems for assigning prominence are conflated in a single group it will appear more prominent in an independent than in a dependent clause. That is, the logical and textual systems offer each other greater choices for assigning prominence. Consequently, logical status and sequence will also be a part of the analysis, resulting in, for instance, a distinction between a Rheme that is final (including final in a simple clause) in an Independent clause from a Rheme that is final in a dependent clause.

3.7 Analytical Model

This section outlines the choices that will be made in the analysis of these texts, detailing the choices in Theme and related choices in the Logical status of each clause.

3.7.1 Theme-Rheme

The pertinent aspects of Theme from the preceding discussion can be summarised as points 1-17. (The relevant discussion is indicated by section number in parenthesis.)

1. In SFL theory, Theme is realised in an English clause by sequence (3.1).
2. Theme is realised in all metafunctions (Ideational, Interpersonal and Textual) (3.1.1).
3. Theme is part of the Textual metafunction and consequently its 'aboutness' refers only to the status of the clause-as-message, rather than its ideational content (3.1.2).
4. Theme operates at the ranks of group, clause and text, but this study will analyse the thematic status of groups within the clause and clause complex (3.1.2).
5. The extent of Theme is limited by the dynamic interaction of subject and finite in all Moods (3.1.2).
6. Themes will be analysed in all clause-like structures (ranking and embedded , finite and non-finite). The analysis will also recognise the logical status (hypotactic or paratactic), as well as sequence (progressive or regressive sequence) (3.1.2, 3.6).
7. At the level of clause complex, there may be a difference between the thematic potential of the dependent clause in a progressive or a regressive sequence (3.1.2, 3.6).
8. More than one experiential Theme can be included before the finite. The Theme that corresponds with the obligatory Subject will be labelled Topical, and any others Experiential (3.2).
9. Where a Theme contains more than one element, it will be called a Complex Theme (3.2)
10. Themes are not formed by constituent structures; Themes and Rhemes follow a wave-like pattern from most to least Theme-like and from least to most Rheme-like (3.2).
11. Special thematic structures (Thematic equatives, predicated Themes) will be analysed in the same way as all other clauses. (3.3)

12. While realised within the clause, the effects of Theme on a text are best characterised by its patterns across clauses (3.4.2).

13. Method of Development better describes an SFL view of the effects of Theme across clauses than thematic progression, which was formulated for a non-sequentially realised characterisation of Theme (3.4.2).

14. The wave-like organisation of the Textual metafunction predicts that a written Text in particular genres will typically have a macro-Theme that predicts the hyper-Theme of a section of a text that predicts the likely Themes in adjacent and gapped clauses. As this analysis is carried out at the clause level, it is included with the analysis of logical status (3.4.3).

The system diagram in Fig. 3.6 represents the options in Theme, as described above. A clause can consist of a Theme and Rheme, but based on the above criteria, we can see that the binary distinction between Theme and Rheme is not sufficiently delicate for this study (see 1.2.3). A Simple Theme consists of one choice from (in order of likelihood) Topical, Textual, Experiential and Interpersonal Themes. Any combination of these Themes produces a Complex Theme. Theme choices are simultaneous because no single choice restricts any other choices; a Theme can be characterised as Simple Textual; Simple Topical; Complex, Interpersonal and Experiential; Complex, Textual and Topical; or Complex Textual, Interpersonal, Experiential and Topical; and so on.

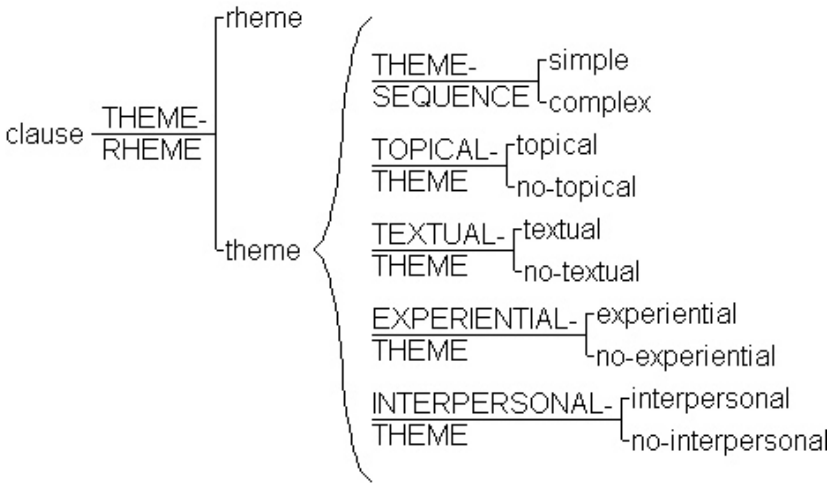


Fig. 3.6 Standard choices for Theme and Rheme

The following paragraph is intended as a short sample of some of the major options in the Fig.3.6. All Themes are underlined.

> The neighbours were very good to her: occasionally some had the children in to meals, occasionally some would do the downstairs work for her, one would mind the baby for a day. But it was a great drag, nevertheless. It was not every day the neighbours helped. Then she had nursing of baby and husband, cleaning and cooking, everything to do. She was quite worn out, but she did what was wanted of her. (Lawrence, 1913/1999, p.73)

The first clause contains a Simple Theme, which is Topical. The next two clauses have identical Complex Themes with both an Initial Interpersonal (*occasionally*) and a Non-initial Topical Theme (*some*). The second clause complex also contains a complex Theme, with the Initial Textual Theme *But*, followed by the Non-initial *it* as the Topical Theme. There is one example here of a predicated Theme, starting *It was not every day*. The analysis shows two clauses with two Simple Topical Themes, *It* and *the neighbours*. A possible Textual Theme, *that*, has been omitted here, and another, *and*, omitted before *one* at the start of the final clause in the first clause complex. Omitted items are not analysed, but they show how meaningful choices in producing text result in variation in Theme.

An alternative approach to Theme is to look at the Theme in terms of the constituents that it contains. That is, rather than classifying the type(s) of Theme, we can identify the roles that constituent groups play in Theme. This allows for comparison at the rank of group across different systems. Fig. 3.7 represents analysis of the status of the constituents of Theme. The units of analysis for the system in 3.7 are groups. These are often, but not always, the groups identified for participant analysis in chapter 2. Each group can be analysed for both Theme function and Theme position, resulting in a full description of Theme types in each clause. In this case, the choice for each group will be exclusive – a group will form part of the Textual or Topical region of a Theme. Groups in Themes are classified as either initial or non-initial in complex Themes (i.e. containing more than one thematic element) or the only Theme in a simple Theme. Groups in Themes are also distinguished by function with the choice of Interpersonal, Textual, and Ideational which is either the Topical subject or another Experiential element in the Theme.

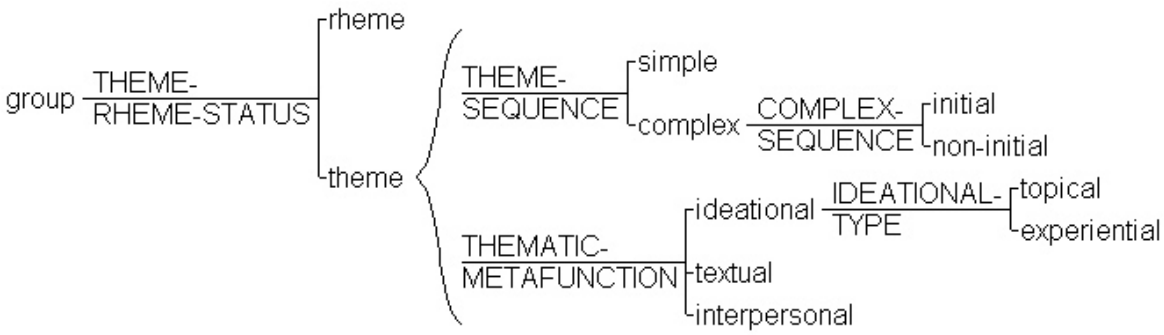


Fig. 3.7 Analytical model of Theme status of groups in clause

In a typical analysis of Theme in a text, the Theme is identified and, where necessary, its separate functions are identified. The aim of the present study is to examine the contribution of textual systems to meaning in the clause. To this end, we are interested in the minimal unit of the group, particularly the (potential) Participant, and to examine how Participant identification and tracking interact with Theme and Information systems. To assist the comparison we will remain at the minimal unit of the group when analysing Theme and Information, although this is clearly not the same unit of analysis for Theme and Information, which can be characterised as 'zones' in the clause. This study takes the group (*e.g.* adverbial, conjunction, participant) as the unit of analysis in all systems and assigns to it a role in thematic and information structure in order to have a common unit of analysis through which to highlight and detail the interaction between the systems. Other elements, such as conjunctions, that are essential to understanding the functions of Theme (this chapter) and Information (chapter 4) in the clause are also analysed. The result is an analysis that details the function of each element in Theme.

3.7.2 Logical Structure

Section 3.6 discussed the interaction between Logical relations and the realisation of some textual meanings, *i.e.* sequence. It is important that progressive and regressive sequences are identified, as the saliency of a Theme or Rheme will be affected by its placement within a dependent or independent clause. These are decisions that are taken at the level of clause, rather than at the group level, so for this network only, the clause is taken as the unit of analysis. Fig. 3.8 outlines the analysis of the logical and hyper or macro status of each clause. Each clause is analysed first according to whether it is a simple clause or clause complex. Clause complexes are then analysed for their logical status of dependency: whether the complex is related hypotactically or paratactically. If the former, the clause is

analysed for dependency and for sequence – Progressive (independent followed by dependent) or Regressive (dependent followed by independent).

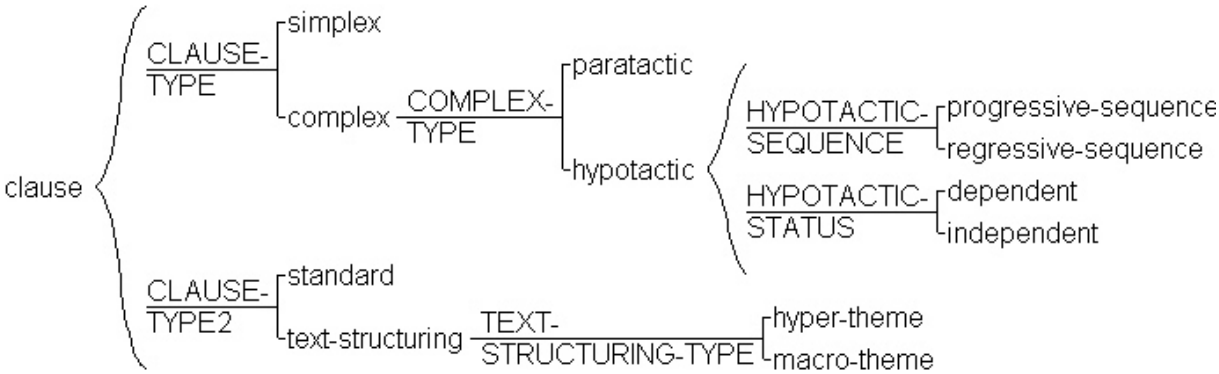


Fig. 3.8 Analytical model of status of clause

Also at the level of clause, and parallel to the logical analysis, is the clause's text-structuring function (see section 3.7.1). The clause as a whole will be analysed for its function of organising larger stretches of text by predicting smaller-scale Themes. Analysis of Method of Development and both hyper- and macro-Themes can then be compared with thematic and, later, information structure analyses.

In the following paragraph, intended as an example of the above scheme, single-clause sentences (Simplex Clause) are shown in increased font size, Regressive sequence Dependent clauses are shown in italics, and Hyper-theme is underlined.

> *In order to understand what you would see* if you were watching a star collapse to form a black hole, *one has to remember that* in the theory of relativity there is no absolute time. Each observer has his own measure of time. The time for someone on a star will be different from that for someone at a distance, because of the gravitational field of the star. Suppose an intrepid astronaut on the surface of the collapsing star, collapsing inward with it, sent a signal every second, according to his watch, to his spaceship orbiting the star. At some time on his watch, say 11.00, the star would shrink to below the critical radius at which the gravitational field becomes so strong that nothing can escape, and his signals would no longer reach the spaceship. *As 11.00 approached,* his companions watching from the spaceship would find the intervals between successive signals from the astronaut getting longer and longer, but this effect would be very small before 10.59.59. [Hawking, 1988 p.96]

The first clause complex previews the remainder of the extract, as a Hyper-theme, in a Regressive sequence in which the dependent clause also has a further clause dependent on it starting with *if*. Following two Simple clauses and a clause complex of a projecting

clause (*Suppose*) and its projected clause, the next clause complex contains two paratactic clauses, starting with *At some time* and *and his signals*. The final clause contains both a Paratactic pair of clauses (*his companions* and *but this effect*) and a Regressive sequence Hypotactic complex, with the dependent clause starting *As 11.00*. Within many of these clauses there are also embedded clauses (*e.g. at which the gravitational field…*). While embedded clauses are analysed for thematic content, they do not warrant separate analysis for logical status as they share the same status as the clause in which they are embedded.

3.8 Thematic Analysis

The analyses in this section mainly utilise quantitative measures which are then interpreted. Also included are other analyses that focus on localised aspects of the sample texts. Various aspects of Theme and Rheme are discussed before combining the results of thematic analyses with results for participant tracking. The aim here is to characterise the thematic choices made in these texts, focusing especially on those choices that contribute to meanings across clauses. The analysis of Theme will be extended from clause grammar into discourse semantics as far as it can in order to assess how much Theme and Participant Tracking and Identification contribute to discoursal choices.

3.8.1 Overview of Theme and Rheme at Clause Level

In total, there are 808 Themes (502 Simple Themes and 305 Initial in Complex Themes) and 808 Rhemes. The same figure is a coincidence and hides the fact that some Themes have no Rheme, and some Rhemes have no Theme, most often because of ellipsis. For instance, in BN, the phrase "*logical link control (LLC) layer"* stands alone as a Rheme in a list which was introduced three clauses previously.

Looking at all six texts at once, Themes and Rhemes can be divided into 2809 groups which are all potentially Theme or Rheme. The remainder of the analysis mainly attempts to describe the tendencies of these groups. 1134 (about 40%) of these groups are placed in Theme and 1675 (60%) in the 806 Rhemes, as illustrated in Fig. 3.9 which shows the proportions of all choices at the group level in the system network for Theme. Of 1134 groups in 808 Themes, just over half (56%) are in Complex Themes (corresponding with an average distribution of 1.41 groups per Theme), with 52% of those offering a non-initial position, suggesting a very small proportion of complex Themes with three or more

elements. (Since every non-initial Theme must have an initial Theme, an equal number shows exactly two elements in Complex Theme. 64% Non-initial to 36% Initial Themes, for example, would reveal an average of three Thematic elements in Complex Themes.)

The choices that appear marked in this register are variations in the metafunctional role played by the different elements in Theme. That is, the unmarked choice for a Thematic element is Topical Theme (just under two in three of all Themes are Topical). A Textual Theme is slightly marked (approximately one in three Themes are Textual), but less so than an Experiential Theme (about one in 8 Themes), while Interpersonal Themes are extremely marked (about one in 140 Themes). In this selection of texts in this register, then, an Interpersonal Theme is a marked choice that is likely to carry greater impact than Textual Themes, which are slightly more marked than Topical Themes.

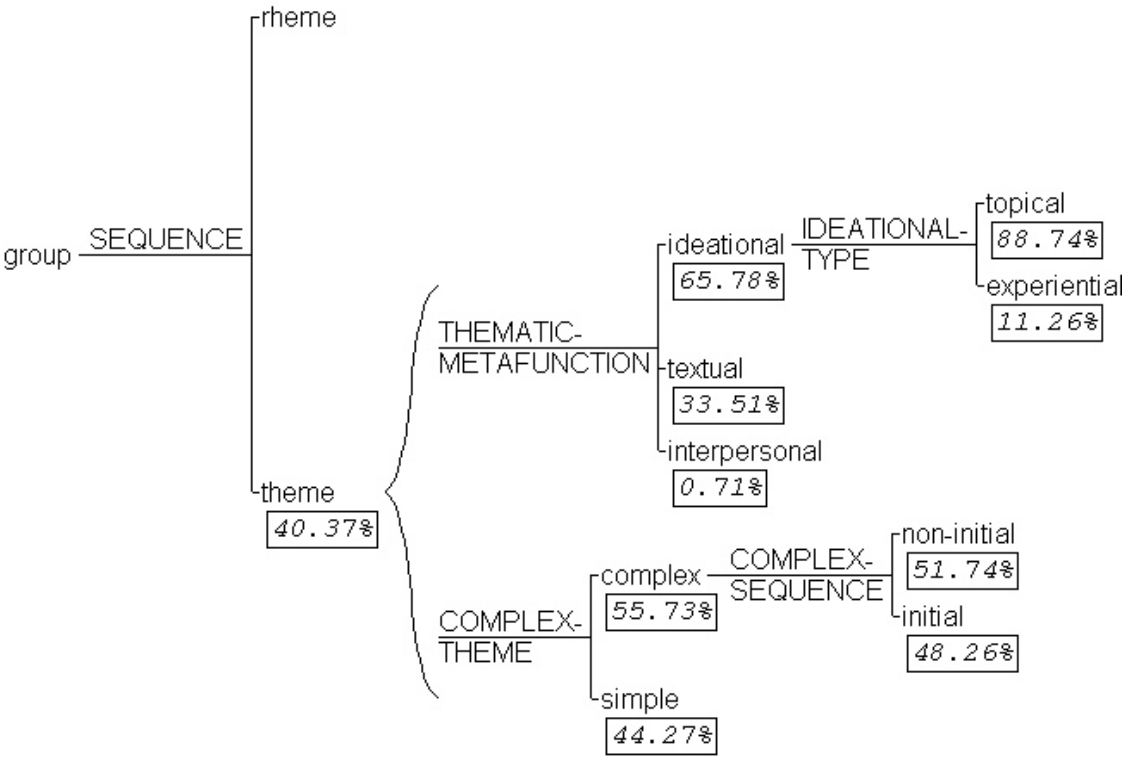


Fig. 3.9 Proportion (%) of choices in Theme-Rheme system for all texts

3.8.2 Theme and Rheme in Individual Texts

In most cases the Theme and Rheme patterns in individual texts follow those outlined above for the corpus of sample texts (see Table 3.2). A few variations in the distribution are outlined here. While the average ratio of groups in Rheme: groups in Theme is 6:4, the range of groups in Rheme:Theme ratios range from approximately 7:3 (for two texts) to

roughly 11:9 (for four texts). The lowest proportion of groups in a clause in Theme position is 29.2% and the highest is 46.1%. This difference is not likely to be the result of author style, as both Text 4 (29.2% of groups in Theme) and Text 5 (43.4% in Theme) are from the same source, sharing the same co-authors, editor and publisher. Since all of these texts have the same discourse function of instruction in technical subjects, the reason for such a wide distribution is probably not to be found in genre. Local discourse features may have some bearing, but this cannot be discovered by looking at Theme and Rheme alone.

	Theme		Rheme	
	n	%	n	%
Text 1: MM	91	34.6	172	65.4
Text 2: EAP	258	43.2	339	56.8
Text 3: BN	51	31.1	113	68.9
Text 4: Arc	42	29.2	102	70.8
Text 5: Sec	89	43.4	116	56.6
Text 6: AN	43	41.8	60	58.2
Text 7: RM	279	39.1	434	60.9
Text 8: FMAM	167	44.8	206	55.2
Text 9: FSI	114	46.1	133	53.9
Total n / Average %	**1134**	**40.4%**	**1675**	**59.6%**

Table 3.2 Distribution of groups in different texts across Theme and Rheme

3.8.3 Simple and Complex Themes: Overview

Theme is divided fairly evenly between 500 groups in Simple Theme (44%) and 641 in Complex Themes (56%). At first glance this appears to be an unusually high proportion of complex Themes, and is probably indicative of the genre. That is, readers of this register should expect to find Complex Themes slightly more often than Simple Themes. Further comparison of texts in this register with texts from other registers would verify this observation. Simple Themes consist of 135 groups in the role of textual Themes and 363 as ideational Themes, of which only two are experiential – imperative verbs in both cases. Table 3.3 provides figures for Themes in all Metafunctions, with subtotals in parentheses. Totals for Ideational Themes are divided, in italics, into Topical and Experiential.

It should be noted that there are 27 groups that have been analysed at two levels of delicacy, producing 2832 groups analysed in total (compared to 2809 distinct groups in

Theme and Rheme). These groups are analysed twice because they play more than one role in Theme-Rheme analysis. This creates an extra 9 Theme groups. In EAP, the clause "*For example, the statement "the copper was removed by the chemical reaction with ferric chloride" is descriptive data.*" is analysed as having both an initial Textual Theme, *For example*, and a Topical Theme, which continues up to *chloride*. This Theme, however, contains a full projected clause with a thematic structure of its own. This is also analysed, producing another Topical Theme "*the copper*". The remainder of the projected clause after *removed* operates as both Rheme in this clause, and as Theme in the larger clause.

		Textual	**Ideational**			**Inter-personal**	Total
			Topical		*Experiential*		
Simple		**138**	*362*	**364**	*2*	0	502
Complex	Initial	(**216**)	(*26*)	**(85)**	(*59*)	(4)	632
	non-initial	(**26**)	(*274*)	**(297)**	(*23*)	(4)	
		242	*300*	**382**	*82*	8	
Total		380	(662)	746	(84)	8	1134

Table 3.3 **Comparison of thematic metafunction and thematic sequence. (Reading across, bold numbers show significant differences to 98% for chi-square test)**

Within Complex Themes, there is a fairly even split between 305 Initial (48%) and 327 Non-initial (52%) Themes – a ratio of 1:1.1, suggesting that about 90% of Complex Themes are comprised of two groups, while only approximately one in 10 contain three or more groups. Five groups are analysed as Simple Themes within Complex Themes, and two are analysed as Complex Themes within Simple Themes. For instance in RM, the Simple Theme "*Many airfields from which fast jet aircraft operate*" includes the embedded Complex Theme "*from which fast jet aircraft*". There is one Textual and one Topical Complex Theme contained within a Simple Theme. Table 3.3 offers figures (in bold when compared across) that show significant differences both between textual and ideational Themes, and, within ideational Themes, between topical and experiential Themes. That is, according to the statistical measures provided by UAM Corpus Tools[3] 2.0 (O'Donnell 2009), in the t-test and chi-square test there are significant differences between these scores ($t=4.4$, $p=0.02$ $\chi^2=(2)\ 22.3$, $p<0.001$). We cannot expect a normal distribution for the nominal data (Butler, 1985) in this study, but the large frequencies allow the t-test to

[3] UAM Corpus Tools (O'Donnell, 2009) has been used for an initial indication of tests of significance for text analysis. However, as the software package offers only $p<0.02$, the scores have been supplemented in the appendices with higher levels of significance

provide a rough guide to significance. However, the data throughout this study is better suited to the chi-square test, which will be used throughout to measure significant differences and associations. Appendix 3.1 Table 1 shows the χ^2 contingency table for comparisons in Table 3.3. As the scores for Interpersonal Themes are below 5% of the total they cannot be included in a calculation. Thus, comparing Textual and Ideational Themes we find that with a highly significant difference between the position of Textual and Ideational Themes (χ^2 =(1) 15.86, $p<0.001$). As χ^2 = (1) 266.38, $p<0.001$, it is safe to posit a highly significant association between Initial position for Complex Textual Themes, and Non-initial position for Complex Ideational Themes (see Appendix 3.1 Table 2).

Thus, there is a clear tendency for Ideational Themes to dominate – both in Simple and in Complex Themes. This is hardly surprising as Topical Themes are obligatory in an unmarked Theme. However, in Complex Themes it is typical for textual Themes to take initial position. To make this contrast clear we can see that 89% of the complex textual Theme groups are in initial position (85 non-initial of 243 Textual Themes), while only 22% of complex ideational Themes are in initial position (297 non-initial of 382 Ideational Themes). In complex Themes, there is most likely to be one element prior to Topical theme, which is twice as likely to be textual than experiential. Within ideational Themes, while simple Themes will invariably be topical rather than experiential, there are clear tendencies for initial complex Themes to be experiential rather than topical (59:26 or proportionally 69%: 31%). Together these figures allow us to characterise these texts as having slightly more complex than simple Themes, with simple Themes typically but not necessarily realised by Topical Theme. Other Simple Themes (totalling 502) will be textual Themes (approximately 27%), such as the coordinators *and, or* and *but* and the subordinators *in order to, that* and *which*, often because they enable ellipsis (see Table 3.4 for all examples and section 3.8.5 for further discussion).

The analysis here reveals a variety of functions prior to the onset of the verbal group in the clause in these texts. It appears that the distinction between Experiential and Ideational Theme, as discussed in Section 3.2, has revealed important differences in Theme. Stopping analysis at the first Ideational element in the clause would have meant that a number of Themes in this analysis would be described as Simple Themes, thereby failing to identify important features in the texture of these texts. The distinction between Experiential and Ideational Theme in this analysis revealed the frequent use of Setting adjuncts in Theme, or of Textual Themes, along with the Ideational Theme, or subject.

That is, the choices taken in the definition and analysis of Theme have revealed various meanings in these texts in the Thematic part of the clauses. As Thompson (2007) points out

> An archive of text analyses is being built up which will eventually allow a more informed comparison of the value of different thematic boundaries – not in order to arrive at a single set of identification 'rules', but in order to understand more fully the usefulness and implications of different methods of identifying theme. (p. 676-677)

The current study is offered as a further approach to the definition and analysis of Theme which, following Thompson, reveals aspects of the Theme zone as a result of the way that it has been defined.

Word	Count	Word	Count
also	2	otherwise	1
and	32	shown in	1
as	5	than	1
as well as	1	that	23
before	1	then	1
but	2	thus	1
by	7	to	17
for	2	to which	1
in order to	2	when	7
or	8	which	18
other than	1	while	1
otherwise	1	who	2
		Total	**137**

Table 3.4 Simple Textual Themes from all texts

3.8.4 Simple and Complex Themes: Individual Texts

The average scores for distribution of features across the Themes of all texts, outlined in the previous section, fail to reveal the great differences across the texts. About two-thirds of Themes in the MM and Arc texts are Simple, in contrast to about two-thirds of Themes being Complex in the AN. Simple and Complex Themes are more evenly distributed across the other two texts, with a tendency to use more Complex than Simple Themes.

The less-than-1% proportion of Interpersonal Themes across all texts is unevenly distributed through the texts. While the majority of texts contain no Interpersonal Themes, suggesting that this is a marked choice in this genre, Sec, RM and EAP contain

all the Interpersonal Themes. Within both texts, these form approximately 1% of Themes, representing a highly marked choice (see Table 3.5).

Text	Simple Theme	Complex Theme	Thematic Metafunction		
			Ideational	Textual	Inter-personal
MM	65.6%	34.4%	82.2%	17.8%	0.0%
Arc	63.6%	36.4%	68.2%	31.8%	0.0%
EAP	47.3%	52.7%	60.4%	38.4%	1.2%
Sec	42.7%	57.3%	78.7%	20.2%	1.1%
RM	40.4%	59.6%	66.4%	32.2%	1.4%
BN	39.2%	60.8%	60.8%	39.2%	0.0%
FSI	38%	62%	62.8%	37.2%	0.0%
FMAM	37.7%	62.3%	66.5%	33.5%	0.0%
AN	28.6%	71.4%	50%	50%	0.0%

Table 3.5 Distribution of Simple and Complex Theme and Thematic Metafunctions across texts

While it is not surprising to find a low proportion of Interpersonal Themes in the genre of engineering textbooks, the variation in the ratio of Ideational to Textual Themes across the texts is noticeable. The AN text bucks the trend for this set of texts, with as many Textual as Ideational Themes. The remaining texts range from 60.8% to 82.2% Ideational Themes.

When we classify Ideational Themes into Experiential and Topical, we find fairly consistent patterns across all texts. Highly consistent, with a standard deviation of 0.05, is the choice within Ideational Theme of including an Experiential Theme. Table 3.6 reveals that EAP is the most extreme text in that it has the least number of Experiential Themes (3.8%), while FSI has the most (16.9%). Typically, Experiential Themes play the Interpersonal role of adjunct, most often separating the start of a clause with a comma. For example the Experiential Theme *On mainframes,* starts the Complex Theme *On mainframes, however, many types of file* (FSI). When Experiential Themes are non-initial Themes, they typically follow a Textual Theme. For example in this independent paratactic clause from RM *and in common with most braking systems, they rely on the principle of energy conversion for their operation* we find, in order, an initial Textual Theme *and* followed by the Experiential Theme *in common with most braking systems* and finally the Topical Theme *they* which exhausts the thematic potential. In some cases, the non-initial Experiential Themes acts as a paraphrase of an initial Experiential or Topical Theme. The first clause in BN is *Chapter 12 (section 12.3.2) introduced the concept of broadcast networks.* The material in

parenthesis is considered an Experiential Theme, because although it may in some cases be equivalent to *Chapter 12* it is not the Topical Theme, of which there can be only one.

Text	Topical	Experiential
Test 1: MM	86.5%	13.5%
Test 2: EAP	96.2%	3.8%
Test 3: BN	87.1%	12.9%
Test 4: Arc	86.7%	13.3%
Test 5: Sec	90%	10%
Test 6: AN	90.5%	10.5%
Test 7: RM	83.3%	16.7%
Test 8: FMAM	88.3%	11.7%
Test 9: FSI	83.1%	16.9%
Average	88.1%	11.9%

Table 3.6 Distribution of Topical and Experiential Thematic Metafunctions across Texts

3.8.5 Method of Development

Most texts in this analysis show typical examples of a consistent Method of Development, particularly at the 'local' level, i.e. over short stretches of text. In the discussion of Method of Development (section 3.4.2), a distinction was made between the approaches that emphasised experiential content and those that analyse non-ideational Themes. This section will start by examining the Ideational Themes to consider their effect on Method of Development, and then look closely at marked Themes to see their effect.

Matthiessen's (1995) suggestion that Topical Themes contribute to the construal of taxonomies in those texts that attempt to describe meronymic relations is certainly borne out in this analysis. For instance, a sequence of typical Topical Themes in MM consists of the following major components:

> The main elements of a typical knee-and-column horizontal milling machine - The elements of a vertical machine - the spindle head - The column and base - Both - The base, - the column - The column - The spindle - The gearbox

Similarly, BN contains the following sequence of taxonomic elements as Topical Themes:

> Comparison between twisted pair and coaxial cable - there - coaxial cable - the cable - it - Both types - coaxial cable - Optical fibre – it

While this is not the only pattern in these and similar texts, appendix 3.3 reveals more than one such sequence in each of these texts, showing it occurs frequently enough to be considered an important generic feature .

AN reveals a sequence of Themes with *We* acting as Topical Theme four times in eight clauses. This text also has 20 Initial Theme groups for the same number of Topical Themes, making its Method of Development dependent on a high proportion of Complex Themes. The RM text shows possibly the clearest and most consistently explicit Method of Development. For instance, the first two clauses start with the marked initial Themes *For many years* and *Over the years*. Other examples later in the text include the marked Theme *On propeller driven aircraft* with the topical Theme *thrust reversal* in a three-clause complex. The Theme in the following clause *Thrust reversal on turbo-jet aircraft* encapsulates and contrasts the previous Themes, by developing previously separate groups in a marked Experiential Theme (Circumstance) into a post-modifier as part of an unmarked Theme.

Table 3.7 categorises initial Themes that are not Topical in a sample of the texts. Clearly the most common type of Initial non-Topical Theme is the subordinating conjunction acting as Textual Theme (53% of non-Topical Themes). After all conjunctions, which enable clause complexing and, often, ellipsis, the next most common non-Topical Theme (15%) is the set of Experiential Themes that function to locate the clause-as-message in textually-construed time and space (e.g. *In Figure 7.1* in Sec, *For many years* and *On propeller driven aircraft* in RM). Other marked Themes are typically Textual Themes that do not fit the traditional category of conjunction, including adverbials (*Finally* in AN) and non-finite verbs (*depending on* in FSI).

Text	Experiential Theme	Textual: coordinating conjunction	Textual: subordinating conjunction	Interpers. theme	Other marked Theme	Tot
Text 1: MM	1	3	6			10
Text 3: BN	1	4	9		4	18
Text 5: Sec	8	6	20		2	36
Text 6: AN	1	6	9		2	18
Text 7: RM	12	17	38		7	74
Total	23	36	82	0	15	156

Table 3.7 Count of Initial Non-Topical Themes in selected texts

3.8.6 Hyper- and Macro-Themes

Although the texts in this study are homogenous, in as much as they are short sections representing extracts from larger chapters, they reveal a variety of hyper- (and macro-) Themes. Hyper-Themes (and Macro-Themes) function as the starting point of a text-as-message. As such, they often state the larger goals of the text (see Table 3.8 for summary).

A number of the texts direct the reader to the immediate and subsequent sections and subsections - alerting readers to the content of the different sections. In BN, for instance, Initial Thematic position is taken up by Macro-Theme groups such as *Chapter 12 (section 12.3.2)*, and *Section 14.2.1*. Other examples include the Hyper-Theme groups *In the model shown* (MM), *In Figure 7.1*, *In Section 7.1.2* (Sec) and *In this section* (Arc). There appear to be no criteria to distinguish macro- and hyper-Themes, so in this analysis the difference was considered to be relative. In terms of the scale of the publication, macro-Themes refer to the organisation of the book as a whole, while hyper-Themes are reserved for sections within the book or chapter.

	Theme	
Text	hyper-	macro-
MM	7	0
EAP	5	0
BN	1	3
Arc	2	0
Sec	10	10
AN	4	13
RM	33	1
FMAM	9	3
FSI	6	0
Total	77	30

Table 3.8 Macro- and Hyper-Theme across all texts

Whole clauses are devoted to orienting the reader to the imminent information to greater or lesser degrees of opacity. The most explicit examples of directing the reader's attention to the organisation of the text are:

2.1 We will begin by talking about the origins and characteristics of the different kinds of noise that afflict electronic circuits. Then we will launch into a discussion of transistor and FET noise, including methods for low-noise

> design with a given signal source, and will present some design examples. After a short discussion of noise in differential and feedback amplifiers, we will conclude with a section proper on grounding and shielding and the elimination of interference and pickup. (AN, p.429)

and

> 2.2 In this section, we shall describe the main architectural models employed in distributed systems - the architectural styles of distributed systems. We build our architectural models around the concepts of process and object introduced in Chapter 1. (Arc, p.31)

Other clauses serve a similar function, predicting what will come next and how it could be organised. For instance the clause *The main elements of a typical knee-and-column horizontal milling machine are shown in Fig. 11.3* (MM) uses an illustration to Anchor (Kong, 2006) the remaining sections in the chapter: all of the selection can be related back to the illustration and this clause. Thus, the diagram and the referring clause function as hyper-Theme for the subsequent text.

Other methods of signalling the organisation of the text include Enumeration, Advance Labelling, Reporting (and evaluation), Recapitulation, Hypotheticality and Question (Tadros, 1985; 1989). A typical Enumeration technique is listing – committing the writer to produce the number of items in the list (Tadros, 1985). Listing is used three times in BN. The first list is expanded immediately, while the second two lists provide the structure for the rest of the chapter. While these predictive techniques may not be considered realisations of hyper-Theme, they perform a similar function.

In addition, all the texts use Titles to organise the text and to provide a starting point for the section as a message. Titles can operate as hyper-Themes, repeating or making the hyper-Theme in the text more obvious. Titles, however, often focus on the Ideational hyper-Themes, and it is often up to the text itself to explain the Textual or Ideational Themes of a section, as described above.

3.8.7 Issues in Theme Analysis: Clause Status and Theme

In the vast majority of cases, it was not difficult to apply the model for Theme-Rheme analysis being proposed in this study to the corpus outlined in Table 1.2. A number of issues were raised in sections 3.1 to 3.5, with implications for the definitions and the choices in this study. The main issue that proved problematic for the model was that of the identification of Theme in dependent clauses, particularly finite clauses. This section

describes, with examples and explanations, where examples in text proved problematic for the analytical model. In each case, the choice taken in the analysis is described.

The analysis of Theme in independent clauses is not unproblematic. However, analysis of dependent clauses, both finite and non-finite, compounds any difficulties in analysis. Halliday and Matthiessen (2004) describe how a "scale of thematic freedom" (p.98) in non-finite clauses operates such that choices in other parts of the grammar may restrict the degree of choice available in thematic position. They also note that these restrictions are ultimately thematic, or Textual, in nature – that is, the restrictions in choice are a result of Textual choices including the choice to combine clauses in a particular way or order. There are a number of examples that were identified in the analysis of these texts that demonstrate greater or lesser degrees of freedom in choice of thematic elements.

In a minority of non-finite, dependent clauses, marked thematic structures closely resemble those of finite clauses. That is, there is an element that is the equivalent of a subject acting as topical Theme. (Since there is no Mood, by definition, in a non-finite clause, there can be no Subject, although there may be an Agent or similar in relation to the process.) In these examples, italicised text indicates an unmarked Topical Theme in a non-finite clause:

2.3 … so that the thrust generated by the propeller is directed forwards instead of rearwards, *the degree of braking assistance* being controlled by use of the engine throttle levers. (RM, section 10)

2.4 This area is either a stretch of rough ground designed to absorb the forward momentum of the aircraft, or a shallow pit filled with pebbles, *both types* being designed to slow the aircraft down on entry. (RM, section 20)

In all such examples, the subject-equivalent element is analysed as topical Theme.

This choice is often not open to non-finite clauses, however. In most cases the Theme zone of a non-finite dependent clause will not include a Subject.

2.5 Many measuring instruments provide more digits than are significant, *leaving* it to the user to determine what is significant. (EAP, p.38)

Extract 3.5 presents a challenge to the analytical model as we must choose between either a theme-less clause, with *leaving* as 'pit' followed by the Rheme, or a Theme realised by the process *leaving*. In this study, the latter option was used, firstly because the relationship of dependency is signalled by the participle form rather than a conjunction, a textually-

motivated choice in itself. More problematic is this unmarked example of a non-finite dependent clause:

2.6 Other forms of noise (e.g. radiofrequency interference and "ground loops") can be reduced or eliminated by a variety of tricks, *including* filtering and careful attention to wiring configuration and parts location. (AN, p.428)

It would be possible to analyse the second (non-finite) clause as being Theme-less: "including" is not a Subject, it allows for very little variation and reveals its dependent relationship in its structure. However, to do so would be to miss on some important aspects of this structure. First, the writer has chosen to make implicit the Logical relationship of Elaboration-Exemplification, (and could be realised by 'which includes'), that exists between this clause and the preceding one. Second, as we saw above, it is possible to add a group that may share features with a Subject before the non-finite verb to give choice in thematic structure. Finally, 'including' could be seen as a form of 'Thematic metaphor', in that it could stand in almost any context in place of 'such as' as a Conjunction to express the Logical relationship of Elaboration-Exemplification. To ignore this aspect of meaning would be to lose an important function of Theme. Consequently, this non-finite clause was analysed as having a Theme realised by the non-finite process.

Other non-finite clauses may be adjoined to the main clause by some form of conjunction. Purpose clauses using the to-infinitive structure, when in initial position, may act as a type of Textual Theme in that they relate the purpose to previous discourse (Thompson 1985). Following the discussion of non-finite thematic choices, initial purpose clauses can be analysed as Theme of the clause complex, as having an independent Theme analysis, or both. In this study, all clauses are analysed independently for Theme, so a Theme analysis must be consistently applied to such clauses as 3.7 and 3.8

2.7 *To overcome* these problems, many small aircraft are fitted with Arrestor hooks (RM, section 21)
2.8 *to prevent* the accidental release of the arrestor hook, a safety device is fitted either on the spring jack, *to stop* the jack from extending, or as a safety bar (RM, section 27)

Since the to-infinitive clause has no Mood, and there are no other groups preceding the verbal group, the most that could be claimed is that the 'to' is a Textual Theme. However, as this structure does not allow the 'to' to be removed from the verbal group it offers no choice in position. This structure allows for other elements to be added before the 'to', *e.g.* "For anyone to overcome these problems…" or "In most aircraft to overcome these

problems…".[4] Although this does not affect the to-infinitive (verbal group) part of the clause, it shows a measure of freedom in terms of what is allowed at the start of the clause. The non-finite to-infinitive verbal group should be analysed as Theme, at least because the analysis then corresponds with other choices based on a dynamic model centred on the process – Theme is identified by the onset of the verbal group. This suggests that this structure is towards the end of the 'scale of freedom' as it allows little choice in the structure.

Our discussion of to-infinitive clauses does not stop at initial position, however. It is frequently difficult to determine whether a to-infinitive clause is part of the preceding verbal group, or whether it could be considered a separate (dependent) clause. This is a significant choice as it affects overall scores and proportions, as well as deciding whether rhematic elements are clause internal or clause final. The following examples highlight this difficulty:

2.9 For many years fixed wing aircraft relied purely on wheel brakes *to retard* their forward movement after landing. (RM, section 1)
2.10 the pull on the tape causes the tape drums and rotors *to revolve* (RM section 22)

In extract 3.9, a separation could be made such that *on wheel brakes* is final position in one clause, with *after landing* final in the next. Alternatively, the clause could be analysed to continue with only *after landing* in final position and the 'to' clause considered a part of the process *rely*. In this and similar cases, the first choice was taken to increase the delicacy of the analysis, because *to* could be replaced by *in order to*, and could be placed at the front. The 'to-infinitive' structure also allows for analysis of the rhematic structure within the down-ranked clause. In 3.10, the same is not true. Since *to* cannot be replaced by *in order to* it is not a non-finite purpose-clause, but belongs to the process 'cause'. Consequently, *revolve* is analysed as part of Rheme.

Other common conjunctive signals that are used to start a dependent clause include *that*, and *by*. Following Halliday and Matthiessen (2004), 'indirect' verbal processes are analysed as hypotactic so *that* may be used as a textual Theme in a finite or non-finite clause of 'verbiage'. In many of these cases, however, there will be further thematic material. For dependent clauses that use *by*, in contrast, none of the examples contained an element that

[4] While the infinitive can be split from the 'to' by a modifier of the verb, *e.g.* "To rapidly/ first/ finally overcome these problems", it cannot be split by another participant or circumstance. Compare: "For anyone to overcome these problems…" with "*To for anyone overcome these problems…".

could be equivalent to subject, or some form of agent. Here we must choose whether the lack of choice in placing *by* in initial position prevents it from being a theme. In this and similar examples

2.11 Instrumentation extends the human senses *by* allowing a numerical value or values to be associated with the measurand. (EAP, p.37)

the logical relationship of Enhancement:Manner is realised by the use of *by*. Without it, the relationship is not clear, and so this was analysed as a textual Theme, which exhausts the thematic potential of this non-finite clause. There are a total of 128 instances of Simple-Textual Theme.

The relative pronoun presents a challenge as it operates simultaneously as textual and topical Theme (Halliday and Matthiessen, 2004, p.85), as in the following:

2.12 The crash barrier (Fig 7) is designed to stop a small or medium sized turbo-jet aircraft *which*, because of some emergency, *is about to overrun* the end of the runway (RM, section 31, emphasis added to all examples)

It is very important to analyse thematic structure for non-defining relative clauses such as these, since failing to do so would leave the Theme-Rheme structure of the clause incomplete, and may neglect aspects of the dynamic interaction between Theme, Participant and Information systems at the heart of this investigation. In this study, relative pronouns were analysed as textual Themes, because it is assumed that there is a textual motivation for deciding to subordinate a clause rather than start a new independent clause with its own Topical Theme.

3.9 Comparison of Theme and Participant Analyses

The Theme zone encompasses a range of grammatical functions, categories and classes; within Theme we find conjunctions, adjuncts, nominal and adverbial groups, and prepositional phrases. So far the analysis has distinguished these classes only when they pertain to Thematic function. However, there is one grammatical class that will take an important position in this study. Chapter 2 examined the textual resource of Participant Identification and Tracking. The interaction between the textual systems of Theme and Participants will now be examined. Appendices 3.2-3.5 show samples of typical analyses of the interaction between Participant Identification, Participant Tracking and the functions of Theme and Rheme.

3.9.1 Participants and Groups in Theme

Participants are not the only groups to appear in Theme and so they must be distinguished. The results so far have referred to all types of group, but as the following discussion shows, the difference is important.

Within the Theme zone there are, in total, 1134 groups. In a participant analysis, there are 562 participants in a Theme position and 395 non-participants. The difference between the two figures of 177 groups is accounted for by the difference in how the groups are identified. In some cases, non-participants may be combined into more than one group in a Theme analysis. For example, the clause-complex in Fig. 3.10 shows the difference in segmentation of units of analysis between the participant/non-participant analysis, where non-participants tend to be combined as a single non-participant unit, and the Theme analysis where different units may be separated if they are relevant to the analysis. In this case there are 8 units in the Theme-Rheme analysis but only 6 units in the Participant analysis for the same stretch of text.

Text	*However*	*it*	*is not possible to have*	*an accurate instrument*	*unless*
Theme	**Theme**	**Theme**		**Rheme**	**Theme**
Participant	Non-Participant		Non-Participant	Participant	Non-Participant

Text	*it*	*is also*	*precise*
Theme	**Theme**		**Rheme**
Participant	Participant	Non-Participant	

Fig. 3.10 Units of Analysis in Theme-Rheme and Participant Identification analyses

The distribution of participants and non-participants across Theme reveals some interesting, significant patterns. Not surprisingly, UAM CorpusTools (O'Donnell, 2009) reveals Ideational Themes to have a highly significant association with Participants χ^2 = (1) 587.7, $p<0.02$ (see also section 3.9.2). (In fact this χ^2 score is highly significant with $p<0.001$.) Conversely, and to a very high level of significance, (χ^2 = (1) 578.6, $p<0.001$), Textual Themes are strongly associated with non-participants (See Appendix 3.1 Table 3). Typical Textual Themes that include a participant are the relative pronouns *which* and *that* when they simultaneously function to combine clauses but also allow participants to be tracked. Other examples include *For this reason* (from EAP) or *As an initial example* (Sec).

There is also a significant association of Non-initial position with Participants and of Initial position with non-participants in Complex themes (χ^2 =(1) 171.7, p<0.001) (See Appendix 3.1 Table 4). This is probably a consequence of the tendency for Textual Themes to precede Ideational Themes and, as described above, for Textual Themes to be realised by non-participants.

3.9.2 Quantitative Analysis of Theme and Participants

To discover how Theme and Participant Tracking were typically associated, the statistical test of chi-square was applied to selected individual texts and to the set of texts as a whole. Presenting and Presuming Reference were compared with position in Theme or Rheme (see appendix 3.1 Tables 5-11 for scores and calculations). The results, summarised in Table 3.9 below, demonstrate a clear tendency for Theme to correlate with Presumed and for Rheme to correlate with Presenting.

Text	χ^2 (all at 1df)	p
MM	12.64	<0.001
BN	7.96	<0.01
Sec	2.47	<0.2
AN	0.68	>0.2
RM	16.2	<0.001
These Texts	32.0	<0.001
All Texts	43.92	<0.001

Table 3.9 **Summary of chi-square analysis of interaction between Theme and Participant Tracking in selected texts and whole corpus**

Of the texts selected for individual analysis, two revealed an association between the system of Theme and the system of participant Identification at the very highly significant level of p<0.001. One text showed a significant association at the p<0.01 level of significance, while two showed no significant differences between the system of Theme and Presuming or Rheme and Presenting.

Looking in more detail at the type of reference used in the corpus in general, we note that there are also significant patterns in the use of pronouns and directed pronouns.

Comparing the use of nominals and pronominals in Theme position, we find that 21.8% of 560 presumed Themes are pronominal, while 92.9% of 637 presumed Rhemes are nominal. UAMTools (O'Donnell, 2007) calculates that this gives a chi-square score of 53.8, ($p<0.02$ at 1d.f. This score is, in fact, significant at $p<0.001$) That is, although pronouns are less frequent, they are typically associated with Theme position (see Appendix 3.1 Table 12). Also significant, within pronouns, directed presumed pronouns (see section 2.4.1.4) are more likely to be in Theme position than Rheme. A chi-square score of 22.5 at 1d.f. ($p<0.001$) reveals that although there are fewer directed pronouns (51 in total) than undirected (981), directed pronouns are more typically associated with Themes (8.7%) than with Rhemes (2.2%) (see Appendix 3.1 Table 13).

On aggregate, possibly as a consequence of the size of the different texts, the results for all the texts combined show a highly significant association between Theme and Participant Tracking, with less than a 0.1% probability that the results are due to chance (see Table 3.9). This suggests that this association is a significant feature of the Textual Metafunction in this genre. Analysis of further texts should be carried out to establish the validity of these results. The following chapter will look in detail at a selection of the texts to characterise the relationship between Theme and Participants and will also look in more detail at how Presenting and Presuming reference are distributed within the Rheme.

3.9.3 Non-participant Topical Theme

The analyses revealed 52 Topical Themes that were not participants. Non-participant Themes were typically 'existential' *there* (examples 3.13 and 3.14) and 'introductory' *it* (examples 3.15 & 3.16).

2.13 *There* are two RHAG assemblies,
2.14 *There* are several widely used patterns for the allocation of work in a distributed system
2.15 On installations where the arrestor hook has to be raised manually, *it* is vital, for safety reasons, to ensure that the hook is securely 'locked up'
2.16 *it* would be wrong to rely on the cockpit indicator light alone

It could be argued that, particularly considering their ubiquity in clauses that offer an evaluation ostensibly free of individual opinion, the 'introductory' *it*-clause has become fixed and so acts as a form of Thematic Metaphor (see Thompson, 2006). While the analysis here does not rule out that possibility, the detailed thematic analysis in this study provides for both possibilities, since each clause is analysed separately. In Thompson's

(2006) terms, this analysis is 'minimal' as it takes the smallest unit as Theme (and has an additional Theme-Rheme analysis in the embedded clause). A 'maximal' analysis would result in the whole of the *it*-clause becoming Thematic with the material after the conjunctive *that* or *to* being Rheme. Other realisations of non-participant Topical Themes (see Table 3.10) include imperatives, such as *do not change*, *increase*, *consider*, *Remember* or *Recall*, and participle forms functioning in non-finite clauses as a conjunction, including *called*, *designing, providing,* and *using.*

Word	Count		Word	Count
It / it	13		ending with	1
There / there	11		giving	1
that	6		ignore	1
using	4		leaving	1
called	3		providing	1
increase	2		Recall	1
consider	1		Remember	1
count	1		see	1
designing	1		starting with	1
do not change	1		**Total**	**52**

Table 3.10 Non-participant Topical Themes

3.9.4 Theme and Context of Culture

This section concludes with a few examples of some marked choices that combine Theme with reference to the Context of Culture, i.e. they make assumptions about the reader's knowledge of participants that are not derived from the participants already presented, and possibly tracked, in the text.

The highlighted participants in 3.17 to 3.22 are all introduced into the text with Presuming reference. Since this is the first mention of each, and there are no other exophoric clues, they are presumed from the Context-of-Culture. The first three examples use Presuming reference to introduce participants in Theme, and the remaining examples all include Presuming reference in the Rheme. In all cases it is very easy for the reader to follow the writer's lead and *presume* that they have already met these participants in the discourse. This releases the writer from the obligation to explain or define any of these terms. In 3.17 the writer presumes that we all know about "the bucket-type doors" on the aircraft,

even though this is the first time they are mentioned. In example 3.18 the writer presumes that there is a ground crew (in Rheme, this time) to do this job – even though this is the first time it is mentioned to the reader. In the writer's context of culture there is always a ground crew.

2.17 During flight *the bucket-type* doors are held in the open position (RM, section 11)
2.18 the hook usually has to be raised manually by *the ground crew* using a lifting rig or mechanical hoist (RM, section 26)
2.19 which forms the tip of *the a/c tail cone* during flight.(RM, section 13)
2.20 The jaws are locked around the roller on the upper surface of the hook by *the roller* operating the jaws (RM, section 24)
2.21 *The need to protect the integrity and privacy of information and other resources belonging to individuals and organizations* is pervasive (Sec, p.252)
2.22 *The provision of mechanisms for the protection of data and other computer-based resources and for securing networked transactions* is… (Sec, p.252)

In example 3.21 Theme is used in a powerfully rhetorical way that allows the writer to force his presumption on the reader. The clause treats the presumption that "*The need to protect the integrity and privacy of information and other resources belonging to individuals*" as 'given' through a combination of Theme position with nominalisation that leaves almost no room for disagreement. The message proceeds from this proposition, and does not allow for negotiation since a grammatical metaphor has transformed the modal aspect of the clause into a noun (Martin, 1991; Halliday and Martin, 1993).

Extracts 3.17-3.22 reveal presumptions of a context of culture that is probably not shared by the consumers of these texts - novices trying to learn about a new field of study. It is essential that both textbook writers and readers are made aware of how such systems operate to exclude (Lassen, 2004). Placing Presuming participants that have no endophoric location, (i.e. they cannot be tracked in the text), in the Theme of a clause reduces the reader's opportunities to negotiate meaning in the clause. The participant becomes 'given' in the sense that it is treated as part of the context. Unless writers are careful, it is far too easy to introduce participants as presumed because they are phorically located in the context of culture rather than in the co-text. Similarly, readers may not realise that incomprehension is the result of the writer taking for granted a participant that is in many ways new to the reader.

3.10 Developments for Theme-Rheme

In an SFL analysis, the Theme-Rheme structure is dominated by Theme. Theme is the meaningful choice considered to be significant in the Method of Development of a text. However, as has been noted, Theme is only the starting point of a clause and Method of Development is only a collection of starting points – there are no destinations in any of these descriptions. Matthiessen (1995) refers to the Rheme as being the location of experiential development through a text. However, there have been few attempts to characterise the patterns that can be expected in Rheme, particularly in spoken discourse.

Further work may need to be done so that descriptive tools for Rheme will be at least as accurate as those developed for Theme. That is, SFL theory needs to be more adequate in its description of Rheme, just as the many descriptions that cover Topic-Comment structures need to be more explicit about Comment (*e.g.* Jacobs, 2001). As with most SFL descriptions of Rheme, there is no attempt to distinguish the different parts of a Comment, giving the whole of the comment a "predicate focus" or the whole of a sentence a "sentence focus." (Lambrecht, 1994; van Valin and LaPolla, 1997). This does not mean, however, that SFL theory needs to tie itself in logical knots, as required by Lambrecht (1994 p.30) when attempting to separate a presuppositional-based definition of Topic-Comment from the predicate-propositional logic of generative grammar to produce a description of intonation in discourse. Rather, an approach that attempts to map how ideational points grow through discourse (Matthiessen, 1995) to match the patterns identified in Theme may help us to understand the role of Rheme as more than merely a residue of Theme analysis. This is particularly important when clauses can consist of a Theme only or a Rheme only (see 3.8.1), making essential a definition of Rheme that is independent of Theme.

3.11 Implications

The texts used in this study, it must be assumed, represent acceptable written English since they are the product of various drafts and the editorial process required by modern publishers. We will assume that, with perhaps a few minor exceptions for 'typos', the language is accurate or 'correct'. Since in many cases, such as those described in the previous section, participants that have not been previously mentioned in the text are presumed, and often placed in thematic position in the clause suggesting that they are

taken for granted, it can only be the case that the interaction of Theme and Reference is a meaningful choice made by the speaker (or writer) within a particular (social) context. That is, this study supports the claim by Fries and Martin, contra Firbas and Clark, that reference is a meaningful choice rather than a form demanded by the co-text.

We have seen that the combination of Theme and Presuming reference, as well as Rheme and Presenting, is no more than the unmarked option; it is by no means obligatory. Disassociating Theme and Reference takes us one step closer to the notion of three independent textual systems operating simultaneously in the clause, combining in marked and unmarked patterns to create meaning. Thus labels such as 'background' for Theme are likely to confuse the issue of the starting point of a clause and what is less salient. Marked options, where Presented referents are placed in Theme, allow for a degree of 'foregrounding' in the position (and syntactic role) reserved for the 'background'. The opposite will also be true.

3.12 Conclusion

This chapter has investigated the Textual resource of Theme. It has discussed a range of definitions for Theme, and looked at possible limits of the 'Theme Zone'. It has defined Theme as the start of the 'clause-as-message' and extended its limit to the start of the verbal process since, from a dynamic perspective, this is where it becomes clear where to locate the obligatory Topical Theme. A discussion of Method of Development demanded an approach derived from Systemic Functional principles that accounted for consistent patterns across clauses, and the notions of macro- and hyper-Theme were used to describe situations where the text predicts its own pattern of development.

A range of texts were analysed and their results discussed. This analysis was then combined with the systems of Participant Identification and Tracking from the previous chapter. We see a clear unmarked correlation between Theme and Presuming reference and between Rheme and Presenting reference, providing evidence for Halliday's original distinction between a co- or con-textually 'Given' element and the starting point of a clause-as-message. Although the unmarked pattern is for Theme to be realised as a presumed participant, participants are often presented in Theme position and tracked participants frequently appear in Rheme. Here is an example which exemplifies this distinction:

2.23 The table provides the surface upon which all workpieces and workholding equipment are located and clamped. *A series of tee slots is provided for this purpose.*

While the first sentence follows the unmarked order with a presumed Theme and a presented participant in Rheme, the second sentence is a marked example in which all participants reverse the unmarked coupling of Presuming Theme and Presenting Rheme. In fact, the clause-final presuming "this purpose" is clearly intended to encapsulate more than one meaning (participants and processes) from the previous clause, and so could be seen as hyper-Rheme (see section 4.5.4).

We have now discussed two of the three main systems in the Textual metafunction that develop discourse semantics within the clause. What we find is that with the systems of participant Identification and Tracking and Theme, we can account for a wide range of textual features. For instance we can explain how participants are introduced into discourse using the resources of Reference. This assigns to participants in text the status of 'given' in the sense that they can be presumed from the context. We also see how the start of each clause provides the framework in which ideational development takes place. These systems cover a wide range of the choices within a clause that are related to the textual metafunction. None of the choices so far, however, account for what happens in the second part of the clause, and we do not have the written corollary of spoken intonation to produce patterns previously described as information structure. This is the subject of the next chapter.

Chapter 4 Information Structure

4.0 Introduction

The continuous stream of discourse is divided, or literally *structured*, into messages of manageable units. In Systemic Functional Linguistics (SFL), these units are called Information (Halliday, 1967a p.200). Within a unit of Information there is an obligatory New and an optional Given. While New is identifiable in spoken English and has realisation rules, Given is both indeterminate and gradable and is recognised as the 'residue' of New – what New is not (Halliday, 1967a p.204). New information has come to be defined as the section of the message that the speaker wants the listener to focus on or attend to; it is *Newsworthy* (Fries, 2002 p.121). 'Newsworthiness' – the focus of New information – can be considered a non-arbitrary, or natural, consequence of the realisation rules of spoken New information: it is generally the loudest, most stressed part of the clause as a result of carrying the tonic foot (the location of pitch change in the message) and so is the easiest to hear (Halliday and Matthiessen, 2004, p.89). This chapter discusses Information Structure, Information Units, and New information, paying particular attention to the differences between information structure and the other systems in the Textual Metafunction, as well as the role of information structure in written English.

We have seen that in the textual metafunction, Theme functions to frame the clause-as-message, while participant identification and tracking serve to bring participants into the text and track the location of their identity. Both of these functions are quite distinct from the characterisation of information above, and so information structure is hypothesised to stand independent of other systems in the textual metafunction in SFL theory. This chapter details the development of information structure and contrasts its functions and realisations with other aspects of the textual metafunction, particularly with Reference in section 4.1 and with Theme in section 4.2, both concepts which other theories have associated with Information Structure. The discussion below highlights opportunities from early formulations in Halliday's descriptions of Information that have been used by other theories to conflate or combine information structure with other systems in the textual metafunction. Also discussed are developments in SFL that have not been incorporated into other theories but that distinguish more clearly the contribution of information structure to the textual metafunction. The constraints that are placed on the

function of information structure in written text are also discussed at length in section 4.3, arriving at some provisional conclusions (in 4.4).

A model is developed for identifying Information Structure and New Information based on SFL principles in Section 4.4, and this model is then applied in a discourse analysis in section 4.5. The model of information structure is then combined and compared with the results for Theme and Participant Identification and Tracking in section 4.6 to produce a model which is illustrated for its implications for reading. Further developments that may improve the model are outlined in section 4.7, with final sections for implications (4.8) and conclusions (4.9). The following, and final, chapter attempts to explain why the realisations of information structure in the spoken and written modes have diverged.

4.1 Information Structure and Reference

This section discriminates between Reference, particularly participant Identification and Tracking, and information structure. It pays special attention to the characterisations of information structure, especially from earlier formulations, that have enabled some theorists to conflate the two functions.

4.1.1 Distinguishing Information Structure and Reference

Information structure for Halliday (1994; Halliday and Matthiessen, 2004) has developed into an important independent resource in the Textual Metafunction. It stands alone in the textual metafunction in being realised by a non-lexico-grammatical resource, that of intonation, with one tone unit being equivalent to one Information Unit. The tonic foot, where the shift in tone is located, is prominent and functions to signal the end of New information, or the focus of the Information Unit (Halliday and Matthiessen, 2004, p.89).

While information structure has always been seen as distinct, particularly as it is realised by the tonic foot in spoken English and is thus independent of clause grammar, some descriptions of information structure allow for confusion with Reference. Reference is realised in the wording of a clause, but Halliday (1994; Halliday and Matthiessen, 2004) views Reference as non-grammatical – it operates beyond the clause to develop Cohesion (Halliday and Hasan, 1976). There is, however, a relationship between information structure and Reference, as Halliday notes that most phoric and deictic items will be realised, in one sense, as given:

> There are a number of elements in language that are inherently 'given' in the sense that they are not interpretable except by reference to some previous mention or some feature of the situation (Halliday and Matthiessen, 2004, p.91)

An unmarked relationship correlates phoric elements with non-New, or New with Presenting reference. It is this unmarked relationship that has often been mistaken as the definition of the terms Given and New. However, in spoken English, Given information is not realised by Reference, and so there seems no reason for Reference to realise Given in written English. The function of New information is not the same as Presenting reference: Presenting introduces participants into the discourse, while New information is the part of the message that the speaker or writer wants the listener or reader to focus on. Finally, any phorically marked element may be New information – in most cases for the purpose of contrast – in speech or writing. Put simply, Phoricity is not a defining feature of Given or New information. The following section (4.1.2) traces the roots of this confusion, before discussing how other schools of linguistics have developed the concept of Information Structure in various ways (section 4.1.3). The discussion now turns to the developments that have clarified the function of information structure in SFL.

Matthiessen (1992) discusses the second-order nature of the textual metafunction, and in so doing provides a further method of distinguishing the systems in the textual metafunction, by describing their functional relationship with the other metafunctions. Second-order here indicates that textual meanings are directed internally to the text itself – towards a semiotic reality – rather than externally, either to the environment or to the interaction between people. Consequently the textual metafunction *enables* potential ideational and interpersonal meanings to mean in context: The textual metafunction *instantiates* ideational and interpersonal meaning.

While stressing that the textual metafunction is not a *post-hoc* addition, Matthiessen (1992) notes that there are no structures that are exclusive to the textual metafunction. That is, it draws on the resources offered by the ideational and interpersonal metafunctions, and adds value to them, underscoring the simultaneous nature of realisation across the metafunctions. For instance, the interpersonal metafunction requires intonation to realise, amongst others, speech function. Intonation requires a tonic foot, but the interpersonal metafunction places no demands on its location. It is the textual metafunction that places value on the tonic foot by assigning it as New information. Thus, the textual metafunction

places value on elements of the ideational and interpersonal metafunctions. This value is realised in systems that exploit prominence:

> Textual meaning in general is characterized by variations in prominence. This variation in content prominence is quite analogous to the pressure pulses of the expression system; and variation in news prominence is actually realized by pitch pulses. (Matthiessen, 1992, p.43)

Consequently, through recurring patterns of prominence and non-prominence, textual meanings are recognisable through text as wave patterns, but ultimately depend on ideational and interpersonal systems to provide meanings.

Martin (1992; Martin and Rose, 2002) also attempts to clarify the separate functions of textual systems at the level of Discourse Semantics. Martin (1992) maps interaction patterns of the textual metafunction across different strata onto the systems that realise them, as in Table 4.1. Martin introduces the discoursal term 'Point' as a discoursal accumulation of News in text, and as the counterpart for thematic Method of Development (see section 3.4.2). Apart from the dominance of Chain and String patterning for Discourse Semantics, it is noticeable that only Point is realised in phonology, and that this resource has no other realisation. Within the Textual Metafunction, Martin distinguishes the systems of Reference and Participant Tracking (see chapter 2) from Theme and Method of Development (chapter 3), and from New Information and Point.

Interaction Pattern	Discourse Semantics	Lexico-grammar	Phonology
Cohesive harmony	Chains & strings	Experiential functions	…
Method of Development	Chains & strings	Theme	…
Modal responsibility	Chains & strings	Subject	…
Point	Chains & strings	…	New

Table 4.1 Some principal patterns of interaction across strata (Table 6.8 in Martin 1992, p.393)

4.1.2 The Development of Information Structure: Early Formulations

Having outlined some aspects of the current position of information structure in systemic functional theory, we will now investigate its origins. Information structure resides inside the Information Unit – itself realised by intonation:

> The information unit is what the speaker chooses to encode as a unit of discourse; the decision is a meaningful one, and a text may be structured into such units in any number of ways all other features remaining constant. At the same time the information unit is the point of origin for further options regarding the status of its components: for the selection of point of information focus which indicates what new information is being contributed. (Halliday, 1967a, p.202)

Spoken text can be divided into Information Units, realised in intonation, which can be subdivided into component parts. An information unit is a 'quantum' of information – a unit of discourse independent of other structural features. Spoken language can always be analysed as information units even if it cannot be analysed as grammatical clauses. The continuous stream of Spoken English is thus divided into 'bite-sized chunks' to make it more manageable for the speaker to produce and for the listener to comprehend.

Within the Information Unit not all elements are equal; some elements are assigned prominence by the speaker. Information Units can be subdivided into an obligatory New and optional Given. New information here does not mean referentially New, but the focus of the Information Unit. Information structure is distinct from clausal grammar, and is realised in English in intonation: "the generalization that the information unit consists of an obligatory new element, realized as tonic, optionally preceded by a given element, realized as pretonic" (Halliday, 1967a, p.204). While this order is the unmarked sequence, it is not obligatory, and New may precede Given. It is here that Halliday's description allows for confusion between the systems of Reference and information structure: "Anaphoric items are inherently 'given' in the sense that their interpretation depends on identification within the preceding text." (1967a, p.206). Halliday's description of the main choices for Given following New in terms of anaphora enable non-functional interpretations to associate information structure with Reference. As Halliday (1967a) freely uses the terms Given and recoverable (and later predictable) in relation to information structure when discussing reasons for making unmarked choices, it is unsurprising that they may be interpreted as the same thing.

The descriptions offered for information structure are also open to a variety of interpretations. When Halliday (1967a) maintains that

> Information focus is one kind of emphasis, that whereby the speaker marks out a part (which may be the whole) of a message block as that which he wishes to be interpreted as informative. (p.204)

it is clear that this definition of information structure is both functional, relating specifically to the speaker's choice of what to emphasise as informative, and ill-defined

since *informative* is only a paraphrase of information. It is also clear that Information focus depends on the Information Unit – here, the "message block". However, when Halliday goes on to describe information structure thus

> What is focal is 'new' information; not in the sense that it cannot have been previously mentioned, although it is often the case that it has not been, but in the sense that the speaker presents it as not being recoverable from the preceding discourse. The focal information may be a feature of mood, not of cognitive content, as when the speaker confirms an asserted proposition; but the confirmation is itself still 'new' in the sense intended. (1967a, p.204)

it is not surprising that many subsequent theorists miss the distinguishing feature of speaker choice amongst the mentions of 'previously mentioned', 'recoverable' and 'preceding discourse'. In fact, this definition could also be used in a theory of Presenting reference that allows for speaker choice independent of preceding discourse (see 2.4.1.4), and so allows itself to be integrated into other theories as a description of reference rather than as a description of Information. Although Halliday and others have improved on this description, it is still referenced by other theorists (*e.g.* Clark and Haviland, 1977; Lambrecht, 1994) as the main source of a functional description of information structure, allowing what is ultimately a theory of Reference (see section 4.1.3 for further discussion).

Finally, the following two extensive quotations should underscore the comparison of Halliday's description of what is typical and unmarked with what is possible and defining in information structure. First, Halliday describes the role that so-called function words play in the system of Information, realised in intonation:

> Thus reference and other closed system items will not carry information focus even when final in the information unit unless they carry contrastive information… This is why the rule about the location of the tonic is often formulated as 'the tonic falls on the (accented syllable of the) final lexical item in the tone group', 'lexical item' being understood to exclude closed system items, those which occur as the unique realization of a grammatical feature and thus from one-member classes. (1967a, p.207)

Here we see that closed system items are *typically* unstressed in speech, and so are *not likely* to be the focus of information. However, that does not exclude them from being so, perhaps for reasons of contrast. Unfortunately, however, it is easy to mistake what is typical, for example in the case of referenced items, with what is defining. This quotation makes a clearer case for speaker choice:

> These are options on the part of the speaker, not determined by the textual or situation environment; what is new is in the last resort what the speaker chooses to present as new, and predictions from the discourse have only a high probability of being fulfilled. Nevertheless the structure

> of the information unit does contribute in large measure to the organization of discourse, by providing a framework within which these options are exercised. (Halliday, 1967a, p.211)

That is, it is always possible for the speaker to choose a marked information structure, but its markedness will be defined by the co-text. What is missing from the SFL theory, however, is the motivation – psychological, social or otherwise – behind speaker choice. This is particularly significant because, in SFL theory, speaker choice is synonymous with meaning. The quotation also makes it clear that the Information Unit, not the clause or the system of Reference, is the context for each selection of New Information; all choices of focus take place within the Information Unit.

Another paper from the same era makes a clearer distinction between New information and what is referred to in Chapter 2 as Presenting Reference.

> Thus if the system of information focus structures the information unit into two elements that may be labelled 'given' and 'new', we must make clear (1) that 'given' and 'new' represent the speaker's interpretation of the relation of what is being said to the preceding discourse – the 'given' may, but will not necessarily, be overtly anaphoric; and (2) that in the case of the unmarked focus 'given' is to be interpreted as 'unspecified as regards any relation to the preceding discourse'. (Halliday, 1976, p.177)

Again there is the use of potentially confusing terms, such as 'anaphoric' and 'preceding discourse', but here Given is clearly distinguished as having no status in terms of Reference.

4.1.3 Other perspectives

The direction that Information Structure has taken within SFL has not been reflected in other schools of linguistics which tend to emphasise the grammatically salient system of Reference which, as discussed above, forms only part of the original SFL formulation. It is the emphasis on referential New that has been at the centre of other theorists' wish to quote Halliday's work, particularly those with formal and psycholinguistic perspectives. Clark, Prince, Dehé and Lambrecht generally cite Halliday's work to define information structure as the formal introduction of 'fresh' referents into discourse. Other linguists, including Steedman, have appropriated the term *information structure*, often with results that are clearly opposed to the original definition. Finally, Vallduvi is representative of a tendency to supplant the functional definition of information structure with a syntactic definition. These perspectives will be reviewed here, and the differences to the Systemic Functional approach highlighted.

4.1.3.1 Psycholinguistic approaches

Chafe and others have attempted to provide the psychological perspective to information structure missing from the SFL account, with the almost inevitable result that the presumed psychological factors are often considered to be outside, or more important than, linguistic features. For Chafe (1970), new information is based on a psycholinguistic model that supports the transfer of information from one interactant to another: New "is information he [the speaker] is introducing into the hearer's mind for the first time." (p.210) New contrasts with old information: in a sentence, "Some of it is information which the speaker and hearer already share at the same time the sentence is spoken" (p.211) and "Old information means that the concept is already familiar to the hearer (or at least the speaker assumes this to be the case)" (p.214). Chafe (1970) employs folk and psychological concepts to account for Halliday's realisation rules for old and new information, and attempts to constrain new information within an entirely psychological framework (Chafe, 1974). New is distinguished from old information "based precisely on a speaker's assumptions as to what is in his addressee's consciousness at the time of speech" (1974, p.111). These unfalsifiable assumptions then, in Chafe's view, *determine* choices in intonation, pronominalization and word order, among others.

Developments in Chafe's approach include the notion of psychological accessibility of referents, with the categories of inactive, semi-active and active (Chafe 1991, p.51). These refer to participants that are presumed to be unknown, previously referred to but not in focus, and in-focus, respectively. Although Chafe (1974, p.112) acknowledges Halliday as the source of the terms Given and New, he makes it clear that the relation to consciousness is his own. More recently Chafe (1995) has maintained this psychological approach while still acknowledging Halliday's influence.

Chafe (1970, p.214 footnote) makes it clear that he is combining Halliday's two systems of Theme and information structure into one when discussing new information. Chafe (1976) develops new information into the distinct categories of New and Contrast, attempts to distinguish other functions of the start of the clause, and rejects the need for Theme (based on an "aboutness" definition of Theme) because it combines a range of other functions which need to be distinguished.

Ultimately, as Chafe (1976) concedes, his hypotheses range from those with empirical support to the speculative. They attempt to use language, particularly the phenomenon of information structure, in order to explain the mind. This can be seen as an alternative perspective on language from that taken by Halliday, whose apparent aim is to look at language primarily to understand its social function rather than its psychological origin. Although Chafe has refined his views, these early approaches deserve a lot of attention because they have had a great influence on other theorists (see sections 2.2, 2.3 and 3.3), especially Prince, who in turn has greatly influenced many computational approaches to reference.

The development of the notions of the Given-New contract and bridging (Clark, 1977; Clark and Haviland, 1977; Haviland and Clark, 1974) were critically reviewed in section 2.2, and will not be revisited here. Needless to say, although Clark uses the term Given-New contract, the current study views its descriptions in terms of reference because of its focus on antecedents. The Given-New contract and Chafe's work were major influences on Prince's approach to information structure, particularly on the category of Inferrable. Prince has in turn greatly influenced others, including Centering theory (see section 2.7.1) and Huang's approach to 'enhanced thematic structures' (see section 3.3 and 4.7). The relationship between Prince's (1981) 'taxonomy of Given-New information' will be reviewed here in comparison with an SFL approach.

In deciding how information may be indicated as Given or New, Prince (1981) is interested in the psychological presuppositions made by speakers about the presumed state of knowledge of their listeners: "information packaging in natural language reflects the sender's hypotheses about the receiver's assumptions and beliefs and strategies" (p.224). Accordingly, Prince sees the speaker's assumed familiarity of the listener's familiarity with items in discourse as the gradable scheme in Fig. 4.1. An item may be New, which is either Unused or Brand New, which is divided into Anchored (in the context) or Unanchored. Inferrable items mirror Clark's concept of bridging, and are divided into Non-containing and Containing, *i.e.* part of another item previously mentioned (probably in a meronymic or superordinate relationship). Finally, items may have already been Evoked in the context, either through the situation or the co-text. In all examples offered by Prince (1981, 1992), Assumed Familiarity is realised by a determiner.

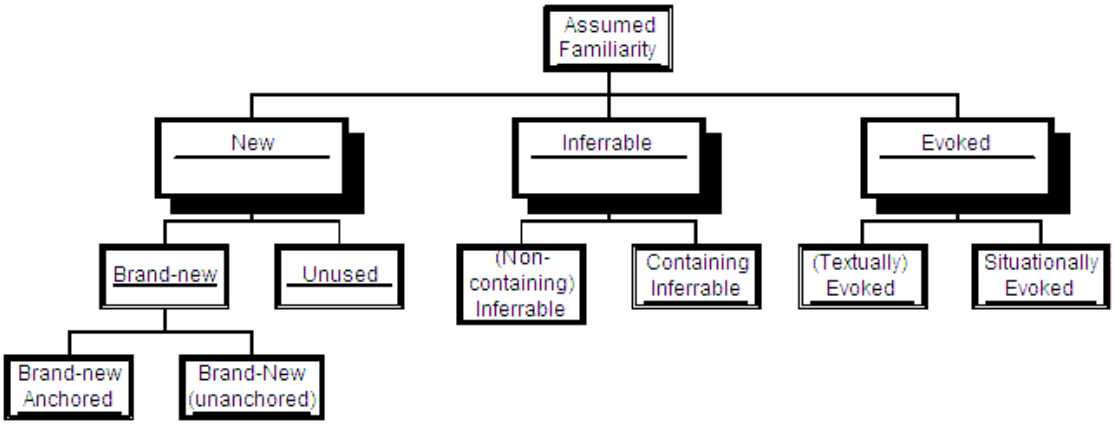


Fig. 4.1 Taxonomy of Given-New Information (from Prince 1981, p.237)

Prince's view of information can be likened to the 'folk' view that ideas are communicated from speaker to listener (*cf.* Maturana and Varela's (1988) bio-cognitive description of communication in section 1.2.4.1). In order to make communication effective, Prince points out that the speaker must take into account the presumed current state of knowledge of the listener. While the SFL approach may benefit from a psycholinguistic perspective, this is not Halliday's intended definition of information – a unit (quantum) of discourse realised by an intonation contour. In fact, Prince patently misrepresents Halliday's information structure when she assumes there may be "more than one marked focus per information unit" (1981 p.227), which by definition is not possible in an SFL approach, although there may be more than one information unit per clause. By replacing the intonation contour with determiners as the realisation of Information, Prince's Given-New taxonomy becomes a theory of reference.

Prince's discussion of 'assumed familiarity' (her preferred term over 'shared knowledge', that rests on no less an unfalsifiable view of cognition) results in a static model of referential knowledge. Entities are "Evoked" by the text or situation, because they have a fixed referent within a lexicon or in the physical surroundings, they are somehow "Inferrable" from a referent in the current discourse, or they are somehow "New" because the speaker says so. That is to say, there is no role here for the listener/reader to construe their own view of a 'reality' – their main role is to match their index of referents with those of the speaker, using the linguistic cues on offer. This privileges the speaker's discourse as the only true discourse and portrays listener-discourse as never more than second best. It also rests on a single representation of reality which must be matched in both speaker and listener in order for understanding to take place. Such a view would be difficult to accept within an SFL approach that relies on a cooperative but independent

construal of discourse (Halliday and Matthiessen, 1999; Thibault, 1999; see section 1.2.4). From an SFL perspective, perhaps the most confusing aspect of the Given-New taxonomy is that we find the term information structure defined by a psycholinguistically motivated version of reference. Based on examples that are offered, all of the categories in Prince's taxonomy, and more, can be accounted for by Presenting and Presuming reference, although even Presenting reference is subsumed under 'assumed familiarity'.

Looking at the results of Prince's (1981) analysis, we find a familiar pattern:

> Nearly all of the subjects are Evoked, whereas less than half of the nonsubjects are. In contrast, one-sixth of the nonsubjects – but none of the subjects – are New. (p.242)

All we need to do here is substitute the words Presuming for Evoked, Presenting for New, Theme for subjects and Rheme for nonsubjects to find the unmarked pairing of Theme with Presuming and Rheme with Presenting identified in chapter 3. It is most unfortunate that Prince (1981) does not offer a more accurate transcription of the analysed dialogue as we may then see where the tonic foot falls, and be able to correlate the New, in Halliday's terms, with the New, in Prince's terms. This would be a valuable comparison.

4.1.3.2 Cognitivist approaches

Approaches that could be loosely described as cognitivist have attempted to incorporate an information structure component into a model of language which combines grammatical, pragmatic and semantic features, but also attempts to explain information structure in terms of cognition – an aspect which is not adequately accounted for in the SFL model (Butler, 2008).

Lambrecht's (1994) approach to information structure can be broadly characterised as combining cognitive, syntactic and pragmatic factors into a model which explains ordering in a sentence. A number of assumptions are made, as a result of the syntactic model within which the study takes place, but are rarely acknowledged. For instance, the primacy of the orthographic sentence is taken as given throughout the volume. That the construct of a sentence may be an invention of literate societies, and not the root of language structure, is not considered (Linell, 2005; see section 5.3). As Lambrecht builds his model of information structure, these assumptions override all other factors. In particular, the

functional explanations that are an integral part of Halliday's linguistic descriptions are generally ignored.

In Lambrecht's (1994) theory of information structure we see that in most cases where Halliday is acknowledged his work is misunderstood, misrepresented or misinterpreted. This is partly due to the poor formulations of theory offered in Halliday (1967a), but is also due to Lambrecht's wish to cite Halliday's functional approach at the expense of clarifications and developments in the theory. Significantly, Lambrecht makes it clear that his definition of information structure is unrelated to focus realised in intonation:

> Finally, I will emphasise that certain prosodic phenomena which have been subsumed under the general rubric "focus" are in fact not related to focus in the sense defined here but to the marking of different ACTIVATION STATES of discourse referents, which in turn serve to indicate certain TOPIC DISCONTINUITIES in the discourse. … I will argue that the overriding purpose of sentence accentuation is not to mark foci but to mark the establishment of RELATIONS between various kinds of denotata and the propositions to which they belong. (1994, p.208)

Here Lambrecht presents the case for a syntactic theory of information based on psychological state, dependent on predicate logic, and incorporating reference. The theory is, further, dependent on a representational model of language – language is designed to represent a reality through denotata (*cf.* section 1.2.4). Yet, Halliday's functional definition of information structure realised by intonation is called into service to provide justification for Lambrecht's syntactic definition:

> Halliday states the following general rule: "The tonic falls … on the last accented syllable of the item under focus." (1967:206) (*sic.*)." (Lambrecht 1994, p.246)

However, this quotation is also misleading. The quotation is not a rule, but the analysis of an example. In context, 'under focus' relates clearly to "the domain of focus in the information unit", and the "tonic falls…" only in the specific example being analysed. At the start of the discussion of which the quoted example is part (see full quote in 4.1.2), Halliday is explicit on the status of any 'general rule':

> This is why the rule about the location of the tonic is often formulated as 'the tonic falls on the (accented syllable of the) final lexical item in the 'tone group'… Stated in this form, the rule is still incomplete, since the tonic may fall anywhere within the tone group; what it specifies is unmarked information focus. (Halliday, 1967a, p.207)

The "rule" is *often formulated* in this way, but it is not a rule; the 'rule' refers only to unmarked options. Lambrecht uses the quotation to adopt a syntactic version of information structure more amenable to his purposes: "A modified version of Halliday's

rule is proposed by Jackendoff, who restates the notion "item under focus" in terms of syntactic phrase structure" (Lambrecht, 1994, p.246). However, this severely distorts Halliday's view that information structure is independent of syntax.

Lambrecht (1994), while referring to the textual metafunction as THEME, notes that

> for Daneš, Halliday and Dik, the formal domain of information structure (functional sentence perspective, theme, pragmatic function) is the sentence or clause. Thus for these authors, as for the author of the present study, information structure belongs to sentence grammar. It is not concerned with the organization of discourse, but with the organization of a sentence within discourse. (p.7)

This misinterprets both Daneš and Halliday, (and possibly Dik) in at least two ways. First, Halliday's definition of information structure distinguishes it clearly from clausal or sentence grammar – an information unit parallels the clause only in the unmarked case. Secondly, an important difference is that while information structure may be *realised* in the clause, its selection and its effects are discoursal (encompassing co-text and contexts of situation and of culture). This is true for Functional Sentence Perspective (Firbas, 1992) as well as SFL. Despite the claim to the contrary, Lambrecht's theory is quite different from the works cited.

Similarly, Lambrecht (1994, p.207) misquotes Halliday thus: "Information focus is one kind of emphasis… the speaker presents it as recoverable from the preceding discourse." (*cf.* full quotation in section 4.1.2: "the speaker presents it as **not** being recoverable from the preceding discourse"). Lambrecht (1994 p.207) continues: "The concept of focus as the element of information in a sentence whereby shared and not-shared knowledge differ from each other is closely related to the one used by Jackendoff." In citing Jackendoff's concept of 'shared and not-shared knowledge', Lambrecht's approach depends on predicate logic and psycholinguistics. Thus, in the space of one paragraph, Halliday's functional intonation-based theory of information structure becomes equated with shared knowledge, syntax and reference. However, that was certainly not the intention of Halliday's paper which, while it does allow for some degree of misinterpretation, has other quotations (see section 4.1.1 and 4.1.2 above) that make it clear that this is not the meaning of information structure. In fact, the paper that includes the previous quotation in Halliday (1967a) states:

> But the non-predictability of the new does not necessarily imply factually new information; the newness may lie in the speech function, or it may be a matter of contrast with what has been said or what might be expected.

> Nor does the specification of new necessarily mean that all else in the information unit is fully derivable. (pp.205-6)

Clearly, New does not mean 'not-shared knowledge', and not-New does not mean 'shared knowledge'. Therefore, to then equate information structure with 'shared knowledge' is a misrepresentation by Lambrecht of the nature of Halliday's New information. Lambrecht appears to be caught between acknowledging Halliday's contribution to the development of information structure and changing Halliday's functional definition into a formal version more amenable to his purposes.

Lambrecht's contribution to Information Structure is taken up by LaPolla (1995) within a Role and Reference Grammar (RRG) framework. While Butler (2005) notes that this results in a more sophisticated distinction than the bi-partite Given-New structure, the discussion above would suggest that in relation to Martin's (1992) system for Participant Identification and Participant Tracking, LaPolla's (1995) network does not account for distinctions between, say *a*, *this* and *a certain* for "unanchored" referents or participants (or Presenting reference: see section 2.4.1.4). That is, Lambrecht and RRG approach Information Structure as a theory of Reference, just as Clark and Prince use the terms Information, New and Given to describe reference. As was discussed in Chapter 2, Martin's SFL system of reference is more comprehensive, and has been shown to be applicable to the texts in this study. This study also proposes that the combination of the systems of Reference, Theme and Information adequately account for the linguistic features discussed by Lambrecht. Although Lambrecht (1994) and LaPolla (1995) attempt to account for the influence of Reference on Intonation, they seem unable to account for instances of clauses that do not contain nominal groups, and so can not be signalled for reference using determiners, or for marked instances in natural spoken discourse that result in alternative meanings (Halliday and Greaves, 2008).

In a basically cognitivist approach, Dehé (2002) attempts to combine Halliday's categories of Theme and information structure into a modified version of Focus and Background. Dehé's (2002) explanation of information structure is based on the speakers' assumptions about hearers' knowledge in context While this approach could compensate for the lack of psychological adequacy in SFL theory, I believe that the theory amounts to little more than conjecture, because it depends on an ultimately unfalsifiable notion of cognitive state:

> I[nformation] S[tructure] reflects the relationship between the speaker's assumptions about the hearer's state of knowledge and consciousness at

> the time of an utterance and the formal structure of the sentence (Dehé 2002, p.104)

That is, Dehé proposes a model that requires an analyst to access the state of knowledge and consciousness of the speaker's presumed knowledge and consciousness of the listener, in order to verify the model. Since Dehé is unable to provide evidence of what these consciousnesses may contain, the model is unfalsifiable.

Dehé appears to group Halliday by implication with other theorists whose view of information structure is based on reference and can be directly related to a cognitivist model of context:

> Subtypes of IS are for instance the *Focus-Background-Structure* (FBS), the *Topic-Comment-Structure* (TCS) and the *Theme-Rheme-Structure* (cf. Halliday 1967b; Jacobs 1992; Steube 1997; Steedman 1991, 2000; Welke 1992 among many others). *Theme* refers to clausal constituents that refer to entities and information known by both the speaker and hearer. (original emphasis p.104)

While the lack of clarity in Halliday (1967a) allows for this approach, once again later developments in SFL theory have been ignored. Although the term Theme is used here, it clearly relates to reference. The confusion between information structure and Theme-Rheme is discussed in the following section. Dehé credits Halliday with the rules of realisation for information structure, notes that in SFL Theme is the point of departure, and suggests that Halliday's notion of Rheme "refers to constituents that contribute new information about the theme." (p.104) In Dehé's model, however, Theme and Rheme are re-interpreted as referentially new and given into a model of Topic and Focus. This is a valid position to take, but I would argue that this position is not viable within an SFL approach, particularly in later descriptions, where it is clearly an unmarked relationship rather than a definition. Dehé cites Halliday to support her position, and concludes:

> the rheme part of the sentence is at least in part, but not necessarily completely, new information. This is in line with the example given by Steedman (2000: 659) (p.108)

What is most surprising is that this quotation follows a section that makes very clear Halliday's position that New and Rheme have an *unmarked* correspondence; New is not necessarily associated with Rheme. As the quotation states, Dehé takes the same position as Steedman, but Steedman's definition does not correspond with an SFL perspective. As I will suggest later in this chapter, and attempt to explain in Chapter 5, the correlation between Rheme and New that Dehé makes may in fact be consistent, but only in the case of written English. I will also argue that analysis of New should start at a minimal unit of

discourse rather than assuming an undifferentiated Rheme which could extend into a number of groups in the clause.

Steedman (1991) attempts to bring intonation within the boundaries of syntax to achieve what he suggests Jackendoff, Chomsky, and Bolinger have failed to do. Employing Combinatory Categorial Grammar, Steedman (1991) claims that the Prosodic Constituent Condition (a "Combination of two syntactic categories via a syntactic combinatory rule is only allowed if their prosodic categories can also combine" p.279) enables syntax to determine phonology. However, the way that prosodic categories combine is not explained. Steedman concludes

> The structure associated with intonation contour really is ONLY surface structure in this new sense, supplemented by annotations which do no more than indicate the information structural status and intonational tune of surface constituents in the extended combinatory sense of the term. (original emphasis 1991, p.291)

and, in a later paper: "intonation structure and surface structure are simply different aspects of the same derivational structure" (Steedman 2000, p.680) Fortunately for Steedman, none of the examples invented to support his case extend beyond one clause complex. He is thus not required to consider when or why marked forms appear in real data. Further, categorising information structure as being 'only a part of surface structure' does not help to explain: how surface structure in written English can function without intonation; the causes or effects of marked examples in spoken English; or the differences between spoken and written English. Placing the derivational rules for intonation in "deep structure" makes them inaccessible, metaphysical and unfalsifiable. Finally, Steedman's main objections to Halliday are, once again, based on the assumption that since 1967 there have been no developments in his theory of Theme. Steedman (1991) misunderstands the SFL model by equating Halliday's Theme with Given and Rheme with New. Steedman's understanding of New and Given are based on reference and anaphora. This is why he finds it difficult to understand Halliday's insistence on putting Theme before Rheme in the clause.

Jackendoff's (2002) approach to information structure is to depend on a range of theorists, excluding Halliday, to develop his own description of Topic/Focus features that are realised in prosody. Jackendoff's description is a mixture of Focus, which he equates to Rheme, and psycholinguistic concepts such as shared knowledge related largely to context and reference, resulting in a syntactic model. Although Jackendoff sees intonation

or syntax as options for a language to realise information, he conflates Information with both Presupposition and reference:

> Given that the overall function of communication is to inform the receiver of something new, it stands to reason that every sentence should have a focus. The communicative function of the rest of the sentence, the presupposition, is to link new information to what the hearer already knows (2002, p.411)
> *Once upon a time there was a little girl who lived in a large dark forest* … cannot presume any previous information; the little girl and the forest are both new characters. (2002, p.412)

That is, Reference, presupposition and New information are all equated in the same sentence-based system to realise Focus. Seemingly in contradiction of associating new information with syntax, Jackendoff notes that information is a discoursal-pragmatic phenomenon independent of syntactic constraints:

> The parallel architecture [of syntax and information structure – NM] permits us to separate the semantic/pragmatic property of being focused from the syntactic and phonological devices used to express it. (2002, p.409-410)

Jackendoff equates information structure with the structure of Topic-Comment, subsumed under a theory of common ground. Finally, in Jackendoff's approach, it is possible to speak a sentence with no information at all. Ritualised greetings are given as an example of interactions where no information is exchanged (2002, p.411). What Jackendoff does not describe, however, is how to account for Focus, or intonation patterns, largely, I would suggest, because of a reliance on an unanalysed, pre-theoretical or folk notion of information:

> Information structure is concerned with the role of the sentence in the speaker-hearer interaction—the means by which the speaker intends the sentence to inform the hearer, in the context of previous discourse (2002, p.408).

Clearly, Jackendoff aims to include the influence of context on the sentence with this description, but the definition of information remains at the pre-theoretical level of *informing the hearer*. In sum, it seems that Jackendoff has tried to integrate all of the linguistic, semantic, discoursal and pragmatic features previously associated with information structure, but the result is a less than satisfactory mixture where it is not clear exactly what is assumed to be the definition, motivation or realisation of information structure.

Vallduvi's (1993) review of formal theories of information structure equates Halliday's Given with formal 'ground', New with 'focus' (1993, p.4), Theme with 'topic' and Rheme

with 'comment', despite Halliday's insistence that this incorrectly prioritises ideational meanings as thematic (see 3.1.1). Vallduvi and Engdahl (1996) compare the resources of English, German and Catalan (as well as Turkish, based on data collected by Hoffman, 1994; 1995) to suggest that syntax is the main motivation for assigning focus. Tonality is considered a secondary influence. Although Vallduvi and Engdahl (1996) present a convincing case, it must be remembered that by brushing aside the differences that exist between most formal and functional definitions of information structure, they are effectively unable to dispute the SFL approach. They have ignored the implications of the metafunctional hypothesis, and taken a predicate logic-based approach that fails to account for interpersonal or textual aspects of meaning within the contexts of real discourse. Finally, as we will see in the following chapter, the sentence may not be as robust a unit of analysis as assumed by Vallduvi and others.

In conclusion, psycholinguistic and cognitivist approaches to information structure, exemplified by Prince, Lambrecht, Steedman and Jackendoff, tend to provide a formal syntactic description of intonation by integrating the range of functions and realisations that are being intentionally separated in this study. By combining reference, Topic and Focus with pragmatics, cognitive states and syntax, it becomes very difficult to identify the exact role of information structure in a clause or sentence, With so many factors to account for, information structure becomes almost impossible to predict. Keeping the functions and realisations of Reference, Theme and Information separate gives the SFL model the advantage that they can be analysed for the meaning-making resources that each system contributes to a clause.

4.2 Information Structure and Theme

Attempts to distinguish the functions of Theme and information structure tend to focus on the distinction between the Speaker-orientation of Theme and the listener-orientation of New, between the limited variation of experiential elements in Theme and the dynamic variety of taxonomically-related experiential elements in New, and between the predicting function of Theme, hyper-Theme and macro-Theme and the accumulating function of New, hyper-New and macro-New. These will be discussed in the following sections, and then compared with other perspectives on Theme.

The aspect of Theme most open to misinterpretation, and confusion with information structure, is the use of sequence as the major realisation of Theme and the role played by

sequence in realising New. First position is the realisation of Theme in English – but this is not the only possible realisation in language (see section 3.1.1). Final position corresponds with the unmarked realisation of New. New is realised by the tonic foot, not by final position, in spoken English. Final position appears to be the unmarked realisation of New in spoken and written English – and it is this that has lead to much misunderstanding, and is the inspiration behind this thesis. This issue will be considered in this and the following sections.

4.2.1 Distinguishing Information Structure and Theme

Information structure in Systemic Functional Linguistics traces its roots back, with Theme, to Mathesius and the Prague School notion of Theme, but in Halliday's (1967a) formulation there is a clear distinction between Theme and information structure. For Halliday, while Theme is the starting point of the message, information structure guides the listener's attention and, significantly, in English (Chinese, French (Caffarel, Martin and Matthiessen, 2004) and probably in most languages) Theme and information structure have different realisations.

A significant distinguishing facet of Information Units is that they are not a component of clausal grammar, but are typically concurrent with it:

> The distribution of the discourse into information units … represents a distinct dimension of structural organization, one that is not derivable from other syntactic features. (Halliday, 1967a, p.203)

That is, while the unmarked option is for an Information Unit to be co-extensive with a clause, meaningful choices are derived from the variation between the two. This is an important distinction between information structure and Theme, which is an element of clausal grammar.

It is worth reminding ourselves that in early formulations Halliday makes a distinction between the unmarked alignment of Theme and Given Information and other possibilities:

> The functions 'given' and 'new' are however not the same as those of 'theme' and 'rheme'. The two are independently variable… But there is a relationship between them such that in the unmarked case the focus of information will fall on something other than the theme (1967a, p.205)

However, it is the difference between speaker- and listener-orientation that is perhaps the most telling in Halliday's earlier descriptions – and one that has lasted.

> The difference can perhaps be best summarized by the observation that, while 'given' means 'what you were talking about' (or 'what I was talking about before'), 'theme' means 'what I am talking about (or what I am talking about now'); and as any student of rhetoric knows, the two do not necessarily coincide. (1967a, p.212)

In this quotation it is Given that becomes a part of the distinction, and is equated with preceding discourse. Remembering that Given is optional, and has no necessary relationship to the preceding discourse, a clearer distinction can now be made: Theme is where I want to proceed from now, and New is what I want you to listen to.

Halliday (1985/2004) identifies three differences between Theme and information structure. The first difference is in the realization of Theme by sequence within clausal grammar, and the realization of information structure by intonation outside of other structures, as discussed. The second difference extends the notion of speaker-oriented Theme and listener-oriented Information:

> The Theme is what **I** want to talk about; the Given is what I want **you** to take for granted. It is true that, typically, the two overlap; I am likely to select my Theme from what is already "Given" for you. But the two are independent choices. Both are of course choices made by the speaker; he is the one who decides, not only what it is that will be his own point of departure for the message but also how he wants to organize his message as information for the listener. … The ***focus*** of the information functions as a signal to the listener that this is what's news, what he is expected to attend to. It may be not previously mentioned, or contrary to expectation, or already under attention but needing its status confirmed. (Halliday, 1985/2004, p.223, original emphasis)

Here we see a clear functional distinction between Theme and New. While stressing the movement from a speaker-oriented Theme to a listener focus for information structure, Halliday emphasises here the function of New-as-news; what needs to be attended to. The movement from speaker to listener is again emphasised:

> The essential point is that the two types of prominence differ; and that they differ as speaker to listener. The Theme is speaker-oriented prominence: it is "what I am on about" (grammarians used to call it the psychological subject). The New is hearer-oriented prominence: it is "what I present as news to you". … Each clause is in this sense a kind of gift, one move in an exchange, symbolized by the change of perspective from me to you. (Halliday, 1981/2002, p.240)

While this clarifies the function of information structure and its difference to Theme, it still does not prevent some theorists linking information to Reference. However, it does help to bring out the concept of New-as-news – what is *newsworthy*.

The third distinction between Theme and information is that the dominant pair are Theme and New – not Rheme and Given. Although Given may be equated with Theme in some theories, Halliday (1985/2004) points out the Theme is divided into Theme and not-Theme (named Rheme) while information structure is divided into New and not-New (named Given). For a message to communicate, New is the only requirement, as in the highly efficient "Fire!"

Halliday (1981/2002) aims to show the parallels between clausal grammar and text grammar (echoed in Martin's (1995b) paper which uses the analogy of fractal patterning to highlight the similarities) and the same point is made in Halliday and Matthiessen (2004) when describing the patterning in groups:

> The diminuendo-crescendo pattern we find in the clause is thus also present in the paragraph, and probably other text units as well: a text can justifiably be thought of as a construct of waves within waves. And this nesting of wave-like structures one inside another is characteristic also of lexicogrammatical organization: among the constituents of the clause in English, endocentric word groups (verbal groups and nominal groups) display this same kind of movement from speaker prominence to listener prominence. (p.243)

The main issue to be taken from this quotation in light of the current discussion is the notion of the speaker-to-listener (wave-within-)wave-like structure of textual organization (also Halliday, 1979), which suggests further layering of distinct patterns of thematic and informational features.

Amongst the range of metaphors that have been used to describe the textual metafunction, Matthiessen (1992) pays particular attention to the metaphor of movement, combining it with the wave-like aspect of the textual metafunction. One aspect of this movement is the constant reviewing of ideational content that takes place in a process that Sinclair (1994) refers to as 'encapsulation' (see 2.4.2). Here the movement allows for textual waves to constitute a part of larger waves: "While the clause often distils its own past … it does not predict its own future; …but it does provide material for a number of possible futures." (Matthiessen, 1992, p.61). The textual metafunction in its enabling and contextualising role is inherently dynamic, adjusting constantly to changes in the context and co-text. This is also suggested by the metaphors that Matthiessen (1992) uses:

> A textual wave or pulse, like any movement, is inherently dynamic – a TRANSITION from one state to another. This reflects the dynamic character of textual meaning: what was new becomes given, what was rhematic often becomes thematic, what was non-identifiable becomes identifiable, and so on. These all constitute CHANGES IN TEXTUAL

> STATUS; and they show how the dynamic character of the textual metafunction involved the notion of TEXT or DISCOURSE HISTORY- the past and future of text development as a semiotic journey. (p.60)

Through constant review of the discoursal status of textual elements, the textual metafunction enables logogenetic development, revealing a further significant distinction between Theme and New. New Information functions to expand on the meanings that form the Theme in a typical clause of written English (Matthiessen, 1995). Dynamically, Themes provide the potential ideational growth or development points that are then instantially developed by New information:

> …while Themes present logogenetic growth-points in the instantial ideational system, News present the range of prominent meanings to integrate with these growth points. Ideational meaning are [*sic*.] thus proliferated as New; they represent a kind of fanning-out from thematic nodes (Matthiessen, 1995, p.43)

It is, thus, New information that distinguishes the most important semantic relations that are construed as any text develops. Matthiessen (1995) highlights the complementary relationship of the systems of Theme and Information, specifically Theme and New, in written English. It must be stressed that New here refers only to the prominence assigned to one of the elements of the Information Unit, not to referential newness. The systems of Theme and New do not distinguish between novel, innovative meanings and meanings which have been reiterated innumerable times within the same discourse, discourse community or context of situation or culture – that is the function of the systems of Participant Identification and Tracking. Meaning is instantiated in text using the same thematic and information resources.

Martin (1992) extends Halliday's notion of Theme at clause level by introducing Macro-Theme and Hyper-Theme at discourse level (see section 3.4.3). The analogy is with the traditional rhetorical devices of Thesis Statement and Topic Sentence, respectively, *i.e.*

> macro-Theme: text::
>
> hyper-Theme: paragraph::
>
> Theme: clause (Martin 1992, p.437)

To describe the function of New Information at the level of discourse semantics, Martin introduces the term 'Point' to complement Method of Development, *i.e.*

> Theme: Method of Development::
>
> New: Point.

Martin (1992) further posits a macro- and hyper-New to correspond with macro- and hyper-Theme, which produces the effect of 'sandwiching' the forward and backward

projections in text, as illustrated in Fig.4.2, particularly for abstract re-worked written text typical of an academic genre. Hyper- and macro-News are not repeats of hyper- and macro-Themes, however, because of the different ways in which they project; macro-News "draw together new meanings which had not yet been made." (Martin 1992, p.456) Hyper- and macro-News can summarise, combine, concentrate and encapsulate prior textual meanings. In the textual metafunction, Themes and News pattern in similar ways on different scales, much like a fractal pattern, to produce "Hyper" and "Macro" (Martin, 1995b).

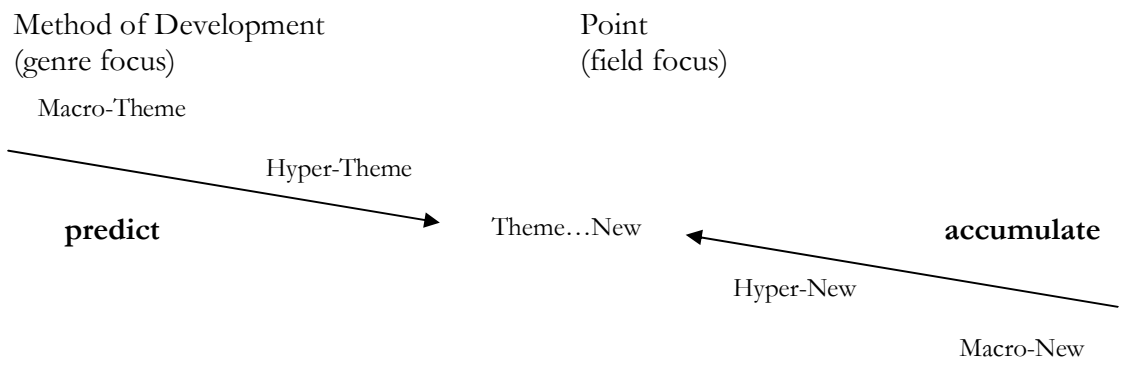


Fig. 4.2 Sandwich texture in abstract written discourse (Fig. 6.12 in Martin, 1992, p.456)

Martin claims that Themes in a clause are both less variable and less frequent than realisations of New information: "Themes angle in on a given field, reflecting a text's genre; News elaborate the field, developing it in experiential terms" (Martin 1992, p.452), reflecting Matthiessen's observation about logogenetic experiential growth. In terms of the movement metaphor noted by Matthiessen (1995), Martin contrasts the restricting action of Theme, as grounding and limiting in terms of range, with the dynamism of clausal New and discoursal News, as open and less predictable: "Where Theme ties text down, point elaborates it, developing it as News." (Martin 1992, p.489)

Thus, in SFL theory we find clear distinctions between Theme and New. Theme is speaker-oriented, the location of potential logogentic growth, it often encapsulates previous meanings, and it is projected by Hyper- and Macro-Theme; Hyper- and Macro-Theme predict textual development. New is listener-oriented, it exploits potential logogentic growth points, generally expands current meanings, and projects to Hyper- and Macro-New; Hyper- and Macro-New accumulate textual developments. Theme is the starting point of the clause-as-a-message. New is what is newsworthy in a clause-as-a-message.

4.2.1.1 Prague School: Communicative Dynamism

Halliday (1967a) acknowledges his debt to the Prague School linguist Mathesius in developing his concept of Theme, which later evolved into the Textual Metafunction. As Fries (1981) pointed out, Halliday and linguists with theories similar to SFL can be characterised as 'separators', because they split the functions of Retrievability and Starting Point of the message, while those that follow Mathesius' original theory could be described as 'combiners' because they conflate the two functions. Probably the most developed version of the Prague School approach of discourse analysis, "Functional Sentence Perspective", is detailed by Firbas. (Halliday (1974) characterises the 'Functional' in FSP as language functioning within the sentence, while characterising the 'Functional' in SFL as language functioning in society.) This section will focus on Firbas' Communicative Dynamism (hereafter CD), and outline the main differences with SFL and other approaches.

The dynamism of Firbas' (1992) CD is a result of the interaction between independent factors. Sequence is one factor. Semantic role (or transitivity role in SFL) and Retrievability are the other main factors. In spoken language, Prosody is an additional factor. Sequence, in Firbas' view, sees CD increase through the sentence, but it is independent of Theme and non-Theme. Firbas' Theme, unlike Halliday's or Mathesius' Theme, is determined solely by Retrievability. Theme is only that which has been mentioned before: "If context-dependent, an element will be thematic. It will be so irrespective of the position it occupies within the sentence" (Firbas, 1987, p.145), and so is similar to Martin's Presuming reference, or Given in Chafe's scheme. Context-independent elements (Presenting reference, or Chafe's New), however, can be either Theme or non-Theme (Firbas 1992, p.72), as shown in Table 4.2.

	Context-dependent	**Context-independent**
Theme	Obligatory	Possible
Non-Theme	Impossible	Possible

Table 4.2 Relationship between context-dependency and Theme in Firbas' Communicative Dynamism

CD divides the non-Theme into Transitional and Rhematic elements. Transitional elements – those involved in the verb phrase in English – are the boundary between Theme and non-Theme and are divided into the Transition Proper and Transition-Proper Oriented Elements. The Transition Proper encompasses Temporal and Modal Exponents (TME). TMEs "provide a link, and at the same time a boundary, between the foundation

and the core" (Firbas, 1986, p.52). Foundation here could be roughly interpreted as Theme and core as Rheme in SFL, with TMEs roughly equivalent to the 'pit' in Theme-Rheme (Hartnett 1995). In English, TMEs are realised by the finite, or by the finite fused with the process. Note that Transition elements can only carry a high degree of CD in the absence of Rhematic elements.

The Rhematic section of the sentence contains any number of Rhemes and the Rheme Proper. Each sentence has only one Rheme Proper, and in this sense could be considered the equivalent of New information in that it is the focus of the clause as a message. The Rheme Proper is the converse of the Firbas' Theme in that it carries the greatest CD in the clause. Fig 4.3 shows which elements in the sentence have the greatest CD, with optional elements in parenthesis. Rheme Proper appears to be defined by what other roles are not: it is not likely to be Transition, Setting, Agent, or non-Final. Although Firbas (1992) claims that there is no direct relationship between Rheme Proper and syntactic function, an analysis of CD must take into account the 'Semantic' factor which tends towards selecting post verb-group participants as Rheme Proper, unless that role is Theme/Presumed and another is Non-Theme/Presented. No SFL Circumstances are analysed as rhematic, as they are defined as 'Setting' and considered to be lower in CD than Rheme, although a sentence-final Setting will carry greater CD than the same setting in initial position because of the contribution of sequence to CD.

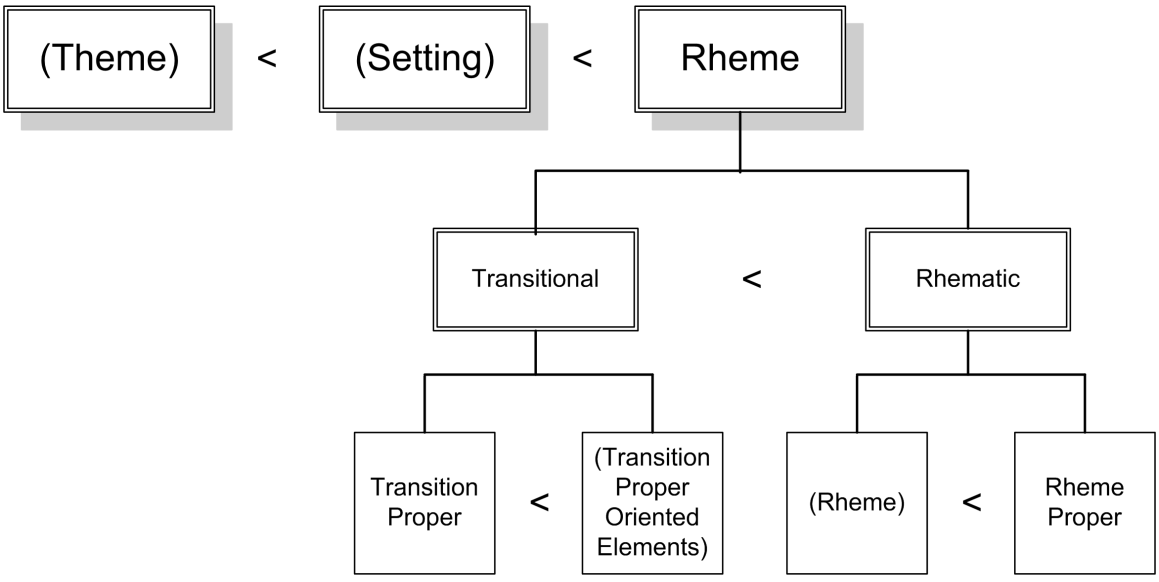


Fig. 4.3 Hierarchy of Communicative Dynamism Showing Proportions of CD and Optional Items (in parentheses) (after Firbas 1992)

For Firbas the co-text alone determines the language used for retrievability:

> The last arbiter of retrievability or irretrievability is not the speaker, but the immediately relevant context, which is binding both for the speaker and the addressee. (1992, p.99).

That is, participant tracking is not a meaningful choice made by the speaker, but is the product of preceding discourse. Halliday's use of sequence to determine Theme is also explicitly rejected by Firbas (1992, p.99), who suggests that characterising Theme in terms of 'aboutness' is the consequence of having identified Theme, rather than a distinguishing feature. Firbas also rejects the notion of the speaker having access to the hearer's consciousness. Contrasting his work with Chafe, whose criterion for givenness "is conditioned by the assumption of the speaker that the information is present in the addressee's consciousness" (1992, p.37), Firbas' criterion "is based on the actual presence of an element in, or its absence from, the immediately relevant context." (1992, p.37) Although the definition of *immediately relevant context* is rarely made clear, Firbas (1986) quotes a study by Svoboda (1981) that suggests a 'retrievability span' of seven clauses, but notes that further research would be needed to clarify such a claim.

Special mention is made of the role of prosody in realising CD in spoken language. Firbas notes, following Vachek, that written and spoken language differ in their mode of realisation, (graphological and phonological respectively), and they also function independently; a reader does not require written language to be spoken, and a listener does not require spoken language to be written down in order to perceive CD. Prosody acts in CD only in spoken language, including the *transposing* of written language into speech when "the transposer must let himself be guided by the outcome of the non-prosodic means." (Firbas 1992, p.221) This, then, gives both the speaker and the transposer of written language the additional resource of 'Accentuation' which can be used to

> reflect or intensify it [Functional Sentence Perspective] in a non-reevaluating or re-evaluating way, but can do so only against the background of perfect correspondence between the non-prosodic CD distribution and the PP distribution (Firbas 1992, p.221)

That is, Accentuation allows a markedness in CD in spoken language that is not afforded to written language, and can only function to focus, or re-focus, on an unexpected element because other factors in CD provide those expectations. The unmarked pattern is the 'perfect correspondence' between prosodic and non-prosodic factors in CD.

As we find ourselves in a quagmire of terms and definitions, some of which appear to correspond closely while others are quite distinct, I will now try to summarise some of the

main differences between Firbas' CD and an SFL approach. Theme for Firbas is almost identical to Presuming reference, but is not confined to participants and so includes referential resources such as ellipsis for other clausal roles. Theme in CD is not sequentially-defined for English or any other language, and 'aboutness' is a consequence rather than a defining feature of Theme. CD uses sequence as an independent factor in the functional perspective of the sentence, and presumes its cumulative effect through the sentence. Unlike SFL, CD recognises grammatical role as significant. In spoken language, including the transposing of written to spoken, prosody adds a further dimension, or factor, to CD and marked patterns are in contrast to the other factors in CD, which are the only resources available to the written mode. It is only this last factor, 'Accentuation', which corresponds closely to Halliday's information structure in spoken English, although all of the factors outlined, and the definition of Functional Sentence Perspective, attempt to identify, as Rheme Proper, the same role as New in SFL. Firbas' CD reveals which element contributes most "to the development of the communication, to which, as it were, it 'pushes the communication forward'" (Firbas, 1972, p.78) The CD model uses different grammatical characteristics to identify the most salient elements in the sentence.

While there are clearly incompatible definitions, such as Theme, between SFL and CD, there are a number of common areas. Ultimately, FSP is attempting to highlight the role that different grammatical structures play in assigning prominence in a clause or sentence, resulting in a single 'Rheme proper' as opposed to the individual textual functions of Rheme, New and Presenting which is what the textual metafunction is also designed to identify. It is probably beneficial to remember the model of CD in order to see how the different systems of assigning prominence might cooperate with each other to produce more salient sections of a sentence.

4.2.1.2 Other Perspectives

There have been further developments outside functional linguistics that focus on the role of sequence in discourse in English. A few of the studies most relevant to the distinction between Theme and information structure will be reviewed here.

As Theme in English is realised by sequence, some studies of the effects of sequence have touched on many of the same issues reviewed above. Arnold, for instance, has approached both reference and sequence from a psycholinguistic and empirical perspective and arrived at conclusions not incompatible with an SFL approach. However,

not all such studies are compatible as they often fail to use linguistic criteria to define categories. Arnold (2001) focuses on Theme, Pronominalisation and Participant Tracking. However, it is difficult to evaluate the relevance of this study as the clearest definition offered for Theme appears to be more related to transitivity role, or the assumption that topical Theme is the only kind of Theme. Birner (1994) describes inversion to be functionally motivated by information structure. That is, inversion allows Given information to be placed sequentially prior to New information in a clause, often in Theme prior to Rheme. Unfortunately, the study uses Prince's (1981) categorisation of referential status for a definition of Given and New, and so would require an extensive review to ascertain whether the same conclusion can be drawn for an SFL definition of information structure.

Arnold *et al.* (2000) look at sentence-building from a psycholinguistic and generative perspective and suggest that, in speech, the demands of real-time processing force both grammatical heaviness and discoursal accessibility to place certain elements in final position in a clause. This is relevant to Theme, Rheme and information structure because final-position in a clause is considered to be the unmarked position for focus in the information unit. It is important to realise, however, that the definition of information structure again follows a psycholinguistic approach and should be equated more with Presenting and Presuming reference. Bearing in mind this terminological confusion, the conclusion may still be pertinent as it suggests that there is a dynamic relationship between the independent factors of syntactic complexity ('heaviness', or simply length of nominal group) and reference, such that a heavier presumed element may take final position rather than a shorter presenting element. This will have an effect on the focus of the Information Unit as presumed elements are less likely to attract the tonic foot. It therefore provides one explanation as to why marked patterns may occur, such that presenting reference may be in non-final position.

4.2.2 The Development of Information Structure: Later Work

While the majority of studies outside SFL assume Halliday (1967a) to be the only source for a theory of information structure, SFL theory has seen a lot of development, typified by the work of Halliday (1994; Halliday and Matthiessen 2004), Martin (1992), Matthiessen (1992), and Fries (2000). The result of this work is a distinction between information and reference, and a focus on the function of 'Newsworthiness'. This section

explores the search for an independent, reliable description of information structure in SFL.

Probably as a result of Fries' (1981) work on Theme and the Method of Development, Halliday (1985) specifically discusses the role of the textual metafunction in relation to written language. However, while Method of Development is discussed, information structure remains distinctly in the realm of spoken text. It was through Martin's (1992) discourse semantics and Halliday and Martin's (1993) analysis of scientific text that information structure in written text became a pressing issue. (See section 4.3.) Halliday (1994) clearly separates information structure from both Theme and Reference, or Cohesion, by allocating them to separate chapters. Halliday (1994) unambiguously identifies information structure with spoken English and considers that it can only be realised by tonicity. The distinction between Given and New is "information that is presented by the speaker as recoverable (Given) or not recoverable (New) to the listener." (Halliday 1994, p.298) The difference in meaning is that New means "Attend to this: this is news." (p.298) while Given means "This is not news." (p.298) Both New and Given are described as choices made independently of any actual occurrence of the element in the con/co-text. Halliday and Matthiessen (2004) place information structure unquestionably within the textual metafunction, and include the resource of information structure in the grammar of clause-as-message, alongside Theme. The definition of Given information remains almost the same, however – that of being presented as retrievable.

We have already seen in chapter 2 that Martin (1992) associates the signalling of retrievability from con-/co-text very strongly with Presenting and Presuming reference. Martin (1992; Fries 2000) also notes that Presenting and Presuming are meaningful choices independent of actual occurrence in text. That is, the system of reference realises retrievability in English. Even though participant tracking only works with participants, other cohesive options, including lexical repetition, activity sequences and conjunction can also function to signal retrievability. This ties the meaning and function of retrievability very much to the system of Reference. Reference works in the same way in both written and spoken English, and so there seems no need to add another system that only appears to work in spoken English, using an entirely different method of realisation, for the same function. That is, it seems highly redundant to develop the system of information, realised by intonation, if it is only to function in the same way as reference. Particularly in terms of functioning in discourse, then, a function is required for information structure that is independent of retrievability. Clearly there is a difference between retrievable and *presented*

as retrievable. This difference, however, is also part of the meaning of Presenting and Presuming reference (Martin, 1992; Fries, 2000).

New may also be distinguished by its discourse semantics. The discourse function for information structure is Point – the corollary of Method of Development (Martin, 1992). In Point, the New of each clause accumulates into Hyper-New and Macro-New which, when combined, provide the listener with the Point of the text. Although this provides us with a different function for discoursal New, it still depends on the same definition of clausal New, glossed by Martin (1992) as "being News to the listener" (p.448) and realised by intonation. We still do not have, at this point, a definition of New that is independent of reference.

Hasan's (1984/1996) description of context and theory of Register (Halliday and Hasan 1985) are explored in depth in another SFL attempt to distinguish the functions of Theme (and Topic) and information structure, especially between written and spoken modes. Leckie-Tarry (1995) utilises Hasan's theory of context to describe Explicit and Implicit texts. Implicit texts are characterised as relying heavily on the context of situation to derive their meaning, and being typically more Interpersonal in the meanings involved. Implicit texts require interactants to derive more meaning from the context; Explicit texts, typically associated with written texts, make their own context more apparent through explicit logical steps and more frequent highly-modified nominalizations. Through the notion of Implicit and Explicit, the modelling of context contributes to the meanings of Theme, Reference and information structure. In this scheme, however, information structure is defined as the degree to which referents are psychologically accessible, and in writing is realised by Reference. Information structure is considered more explicit when "the context of the text can be relied upon to provide clues to the interpretation of information." (p.150) Leckie-Tarry interprets information structure to be almost identical to Presuming reference, noting that Given information can be recoverable

> because it has been mentioned before (i.e., recoverable from the text), because it is something that is in the situation (that is recoverable from the context of situation) or because the speaker expects the listener to have prior knowledge of the entity or phenomenon (i.e., recoverable from the context of culture). (1995, p.151)

Within the current study, this clearly represents phoric options (see section 2.4.2 and appendix 2.2).

In Butler's (2003) comparison of SFL with other structural-functional theories, SFL is evaluated more favourably than both van Valin's Role and Reference Grammar (RRG) and the Functional Grammar (FG) of Dik and his followers in terms of its ability to account for focus in attested marked spoken clauses (p.164-5). Perhaps Butler's (2003; 2005) greatest criticism of Halliday's approach to information structure is given over to Fronek's (1983) discussion of tone units in SFL theory. Fronek (1983) discusses the apparently contradictory statements made in earlier papers that suggest that Given is both anaphoric and non-focal – a 'combining' approach of sorts. However, Fronek's critique fails to explain the syntactic motivations for the irregular ordering of words that he claims should replace Halliday's theory of information structure. In fact very little, apart from a belief in syntax and FSP, is offered in place of the SFL theory, despite an awareness of Quirk's (Quirk *et al* 1972; Quirk and Greenbaum 1973) corpus-based approach that appears to support Halliday's proposals. Collins (1991) also generally dismisses Fronek's (1983) objections.

A more convincing criticism highlights the lack of a psychological dimension in SFL theory, particularly when compared with RRG and FG (Butler, 2008). SFL theory seems to pay little attention to psychological aspects of language production; it is often characterised as a theory of discourse analysis rather than a psycholinguistic theory. However, as discussed in 1.2.4, I believe that this is partly because of a deep divergence between a psychology which serves a view of language as distinct from reality and a psychology that can explain the continual (re-)construal of a language-based understanding of the world (Halliday and Matthiessen, 1999). As its stands, SFL theory does not rely on a psychological description, failing to incorporate concepts such as activated or background. However, such concepts themselves often fail to demonstrate their own validity (see discussion in section 2.2.2). While I agree with Butler that SFL must be able to demonstrate psychological validity for its linguistic theories, I do not agree that current studies that may be of use to SFL should be incorporated uncritically, because of their divergent philosophical approaches. When psycholinguistics is better able to demonstrate the reliability of its theories in respect of neurological processes, I believe that SFL will be forced to adopt a more psychological approach (Moore, 2008b). Section 5.2.1 attempts to integrate a psychological/neurological explanation with the model of information structure proposed in this study.

It is in Fries' work that we probably find the clearest distinction between retrievability, as in Presenting and Presuming reference, and the function of New, in terms of

'Newsworthiness' (Fries 1992; 2000; 2002). Fries (2002) prefers a definition of information structure that does not use Halliday's notion of retrievability as a central element, but prefers "to rephrase the definition of New positively, as 'information which is being presented as "newsworthy"'." (p.121). This definition highlights Halliday's point about the speaker's instruction to "attend to this", as well as making a clear distinction from Participant identification, which Fries describes as causing "an ambiguity, since recoverability is often used to describe the difference between Given and New." (p.122) The distinction between Information and Reference is made clear:

> The opposition between Given and New has no essential relation to reference, while the opposition between presenting and presuming reference concerns referential identity only. It is true that the two concepts generally correlate … However, there are many exceptions to this correlation (Fries, 2002, p.95-96)

The fact that there are so many exceptions demonstrates the independence of the two systems. Evidence for the independence of the two systems is also found in the manner of realisation in the spoken mode – Reference in the grammatical class of determiners and Information in the phonological class of intonation. In most cases, Fries (2000, 2002) follows Halliday by discussing information structure as realised in intonation in speech. He does, however, explicitly discuss information structure in written English. We will return to this in section 4.3.1. Thus, the current study retains this important distinction between reference and newsworthiness.

4.3 Information Structure and Written Text

For many years, particularly in earlier formulations, Halliday made it explicit that information structure is a feature of spoken English. (See Quirk *et al.* (1972) for a similar view). Its realization in intonation forms a major part of the resources in English, independent of clausal grammar, but as intonation is not expressed in written English, most of the seminal papers on information structure only discuss spoken language. As the concept and analysis of information structure developed, discussion of written English increased. The review below outlines some of the major contributions in Systemic Functional Linguistics to information structure in written English.

4.3.1 Theme + Rheme and Given + New in Written English

Although not the first to do so, Martin (1992) directly addresses the issue of information structure in written English. In spoken texts, Given information is indeterminate in the Information Unit, except when it is post-tonic. When identifying New in spoken English, Martin (1992) takes the 'minimal' option:

> In order to simplify the interpretation of New presented here, only the minimal domain of the New will be considered; this will be taken as the highest ranking clause constituent (usually a ranking group or phrase) the tonic syllable falls on the final salient syllable of. (p.451)

This is presumably the same rationale behind Martin's identification of New as the final constituent in a clause in written English:

> In a written text the domain of the New is even less clear than in writing [*sic*. speaking?]. But information structure remains an important aspect of texture in writing, even though intonation is not explicitly realised. Taking New as the final clause constituent … displays the same difference in the range of realisations noted for the spoken text considered above. (1992, p.452)

Martin points out that this heuristic definition of New produces the same range of experiential elements in New position in written text as in spoken text, particularly when contrasted with thematic meanings. This quotation, however, drives to the heart of the problem of information structure in written English: intonation is not realised graphologically, and as New information relies on intonation for its realisation in spoken English, it may not be possible to identify New information in written English. Martin clearly believes it is possible, and provides example analyses using the 'minimal' definition described above. Martin does not, however, address the issue of interaction with Theme. That is, by equating New Information with final position in a clause it is not possible to vary New information in written text in relation to Theme; unlike spoken English, New cannot be found in Theme and Given cannot be found in final position. This returns information structure to Mathesius' definition, against which Halliday reacted when separating Theme and Information, reducing the dynamism created through the tensions between various marked and unmarked options of Theme and New. It could be suggested that in written English special 'thematic' structures (as well as passive voice and other marked features of written English) have developed and become more frequent in written English as a way of manipulating clausal elements into New position. This approach also realigns the SFL approach with Firbas' FSP (see section 4.2.1.1) by assigning to intonation in speech a function independent of written English.

One writer who has possibly done more than any other to bring to our attention the particular issue of information structure in written English is Davies (*e.g.* 1989). Davies (1986) notes the role of Tonality in realising information structure:

> Tonality specifies how the text is structured as a sequence of information units or 'blocks' of meanings by segmenting the text into a sequence of tone groups. All discourse is patterned into 'lengths' of information, as it were, explicitly recognisable by sound in speech, but only suggested in writing. (p.202)

There are, according to Davies, right and wrong ways of assigning tonality and tonicity to writing. That is, there are rules that govern the correct intonation for writing:

> The first kind is the sort of rule which requires the information structure to be compatible with the non-prosodic cohesion; the second kind is simply the correct realisation of an information structure which is correctly compatible with the non-prosodic cohesion in correct intonation. The second kind is less interesting, in that such rules merely echo the normal rules for realising intonation in speech (Davies, 1994b, p.200)

Tonality can be either cohesive and correct or incorrect, acting against the cohesion of the text. Davies spends some time describing how these rules can be identified, starting with punctuation.

In Spoken English, there is no direct correlation between pauses, breaks or gaps and information structure, and "nor is there any direct relationship between the beginnings and ends of information units in speech and the placing of punctuation marks" (Davies, 1986, p.201). Although often considered a guide to both pausing and the intonation patterns of spoken information structure, punctuation offers a precise guide to neither. This begs the question where tonality may be inferred from. Davies suggests that it is the grammar that provides the key:

> Punctuation in writing can be a rough guide to tonality (the 'real' work is done by the underlying grammar) (1986, pp.202-203)
> So although intonation is not physically visible on the page, in that there is no notation for it, it is clear that it is deducible from the grammatical properties of the writing that we can see (1994b, p.203)

For instance, Davies reveals the importance of transitivity roles in providing clues on how to produce an appropriate information structure, suggesting that a reader pauses "not so much because of the comma, but because of the syntactic boundary between the two clauses" (1994a, p.79). That is, the end of a clause will, in the unmarked case, produce a tonic foot to indicate an Information unit, regardless of any punctuation marks. Thus, grammatical structure (including clause length and transitivity), punctuation as a rough

guide, and cohesion (see below) will determine the tonality and information structure produced when reading aloud (Davies, 1994b). However, we are now facing a rather long list of features that all show only a vague correlation with information structure, and we are yet to see which resources produce marked cases.

Information structure has a cohesive effect, and is distinguished from non-prosodic cohesive resources (Davies, 1989). One reason Davies considers information structure to be cohesive is that he often equates it with the theory of Reference similar to Presenting and Presuming and investigated in Chapter 2. Davies (1994a) characterises 'Given' as a referential-based concept closely related to Presuming reference:

> If information at the beginning of the information unit has been mentioned before, or is in some way anaphoric, then it is Given and the status of New information is only assigned to what follows it. (p.82)

Having described as erroneous the matching of reference with information structure by non-SFL theorists, it is worrying to find this kind of suggestion from within SFL theory – although it is also present in Cummings (2000), and is similar to the Quirk *et al.* (1972) definition. Davies proceeds to discuss how Given information is bound to the context of situation and to the context of culture:

> These ties help to create both cohesion and coherence and in so doing create texture independently of and in addition to the ties create (*sic.*) by the non-prosodic resources of the grammar. (1994a, p.85)

Davies (1986) argues that the resources for Martin's system of participant Tracking are prosodic, and act in conjunction with reference. This is clearly not a position that this study could adopt. Having established that the system of participant tracking provides the location where participants are tied to the co-text or contexts of situation or culture through Reference, with all of the instantiations available in that system, it seems unlikely that the unrealised system of intonation in written English will fulfil the same function, notwithstanding the argument that there is likely to be an unmarked correlation between Presuming reference and Given information.

The picture so far seems a little confusing. Davies maintains that information structure is realised in written English in tonality, as in spoken English, even though tonality has no means of expression in written English. Tonality is *only suggested* in written English by punctuation, reference and grammatical structure. However, punctuation offers only a rough guide for tone groups, and can be disturbed by reference. The relationship between grammar and information structure is also far from transparent. Davies' solution is to read

written text aloud, based on the assumption that 'aloud-reading' is no different to any other reading – just louder:

> So if it is the case that when we read we identify the rhythm and intonation of the text, it must be that we only do so because it is worth our while to do so, which in turn must be because we are getting meaning out of it (1994b p.78)

That is, when reading aloud we derive meaning from the text through information structure. Davies maintains that the text contains an inherent rhythm and intonation that are transparent. If we assume, with Davies, that reading aloud brings into existence tone groups that have no verifiable existence on the page, or in the mind, prior to loud reading, then we can consider some readings to be correct and others incorrect because they can be compared against the 'true' test of information structure, *i.e.* tonality, which is revealed upon reading aloud.

In Davies (1989) the importance of tonality and tonicity when reading aloud is exemplified by Shakespearian verse. In another example, a reader is invited to 're-speak' transcribed dialogue. Comparing the 'original' and re-read versions for information structure reveals differences, because in reading aloud the "re-speaker" has, according to Davies, only the non-prosodic cohesion to rely on in order to produce meaningful prosodic patterning. The original speaker, however, was motivated by other aspects of the context. This demonstrates that information structure in spoken language, as realised in intonation patterns, is distinct from the choices made by 'aloud-readers'.

In Davies' (1986) read-aloud protocol, the comparison of a school student's reading of a textbook against a teacher's reading reveals a discrepancy in both the number of information units and instances of New information. According to Davies

> the child SHOULD find many more elements new than the teacher. The teacher … was familiar with the ideas it [the textbook] discussed; the child … was not. (p.210, original emphasis)

In other words, Davies notes the effect of how different contexts of culture resulted in a different patterning of the message for the two readers (see 2.5.2.1). However, this does not correspond with the view that information structure is an inherent property of the text. Davies (1994a) notes:

> Reading aloud requires us not only to give the text AN intonation but also an intonation which works, *i.e.* one which not only realises an information structure but one which is congruent with the other meanings of the text, especially those of Theme and Cohesion. There will often be more than one information structure that is adequate and appropriate (to the rest of

> the structures in that sentence) than others, but the inappropriate ones are far more numerous (p.76)

If two readers can make sense of, and read aloud, the same text in different ways and both are right, then information structure must be a result of the interaction of reader, text, time and audience, rather than the choices made solely by the writer, making it quite distinct from the system for spoken English. In this case, an 'aloud-reader' partly takes on the role of a speaker in deciding where to assign tonality – what to present as *newsworthy* – as a result of the interaction between text and reader. It must be remembered that it is impossible to talk English and therefore to read a text aloud without intonation and therefore without assigning some kind of information structure, but that does not mean that the text contains a particular information structure – only that the reader is able to produce one when forced. We can conclude, then, that the text does not have just one inherent information structure that was intended for the reader to discover, but the meaning provided by information structure in reading is the result of the interaction between writer and reader – whether the reader reads aloud or silently.

Davies (1994b) makes use of transcribed dialogue, lines of verse, and script from a play – *The Importance of Being Earnest* by Oscar Wilde – to discuss viable and incorrect patterns of intonation, but these all represent samples of written English that are meant to be spoken, rather than written English that is meant to be read silently. They are intended to simulate features of spoken English as closely as possible. They can be easily compared with a typical sentence from a random page of academic prose:

> A society's social mode of production is characterised by the social relations which determine the form (or forms) of access to resources and means of production, organize labour processes and determine the division and circulation of products of social labour. (Godelier, 1984, p.104) [5]

This sentence is, almost intentionally, 'a real mouthful'. Reading it aloud is likely to cause strain for the speaker. This is because the passage is not intended to be read aloud, but functions far more effectively as language that stays silent. The sentence functions as written English precisely because it does not require articulation. While there may be a system of information structure, it needs no reading aloud to be realised. If written English has evolved to function without notation for intonation it is probably because such notations are not required to realise information structure or any other function. In other words, if we really needed tonality to be realised in written English, marks would

[5] This text was randomly chosen from a book in a previous field of study – cultural anthropology. The quotation was found on the first page that I opened.

have evolved to realise it. Tonality is a feature of spoken not written language. If information structure is a function of written English it may require no intonation to realise it. Thus although Davies has made great progress in identifying the role of information structure in written English, this study will not use the practice of reading aloud (and so does not follow Banks' (2004) division of New by degrees which also depends on reading the text aloud). This study will work on the premise that what happens in reading need not reflect exactly the processes involved in listening, and that since written language does not realise tonality, tonality plays no part in the realisation of information structure in written English. Section 5.2 examines the consequences of these assumptions in detail.

In *Lexicogrammatical Cartography*, Matthiessen (1995a) makes explicit the differences between the realisation of information structure in written and spoken English, using the term 'Culmination' to describe the tendency to place New at the end of a written clause:

> CULMINATION is the resource for assigning informational prominence in writing in terms of newsworthiness to constituents in the clause. Culminative status is realized by the relative ordering of elements towards the end of the clause. … Culmination complements the thematic status assigned by Theme. Once the local context has been established, the issue is what the main point of information is. (Matthiessen, 1995a, p.600)

Information structure is here characterised as newsworthiness and point and distinguished from the function of Theme. For Matthiessen, the role of intonation in the realisation of information in spoken English is independent of clause grammar:

> … the clause and the information unit come from two simultaneous grammatical rank scales, viz. (i) clause – group/phrase – word – morpheme; and (ii) information unit. The first scale is the same as in the grammar of written language. The second is specific to speech. (Matthiessen, 1995a, p.603)

Matthiessen's information structure in written English performs the same function as in spoken English but, as with Theme, is realised by sequence rather than intonation, which is not realised in written English. It is interesting to note that here, as in Martin's description, the flexibility offered by the separate resource of intonation in spoken language, so that Theme and New or Rheme and Given can be combined, is no longer considered available in written English. In this system, New remains constant in clause-final position, while meaningful choice is achieved by ordering the elements of the clause into final or non-final position through: voice; the movement of adjuncts; ergativity choices in the Process of the clause; the flexibility of position for associated participant roles; and special informational structures (see section 4.2.2 above). It must also be

remembered that there may be a similar culmination from Given to New within units below the clause (Halliday and Matthiessen, 2004).

Fries (1992, 1995a) investigates Theme and information structure in written language. Fries' approach to information structure in written English is similar to Martin's and Matthiessen's in that he sees the end of a clause as the unmarked location of New information. While noting this only as a very strong tendency, Fries also falls back on the loud reading of written English in order to find marked examples. Citing Chafe's correlation of about 60% of intonation units with clauses, Fries (2000) suggests

> Because of the demands of the written language, I believe that this correlation is stronger in the written language than it is in the spoken language. As a result we may use position at the end of the clause as a rough and ready indicator of the location of New information (p.97)

In earlier studies final position in a clause is termed the N-Rheme (Fries 1992; 1994; 1997) to describe the expected position of New information. Fries states this as an explicit hypothesis: "N-rhematic information contains the newsworthy information, information which is in focus in that message." (1997 p.233); only the N-rheme of an independent clause is considered to relate to the discourse goals of a text. One aspect of Fries' analysis that differs from a typical clause-based analysis in SFL is that he analyses information structure in the T-Unit: "an independent clause, together with all hypotactically related clauses and words that are dependent on that independent clause." (1995b, p.49) That is, apart from paratactically related independent clauses, this definition relates to one sentence as the unit of analysis. Thus, there will be as many Themes and as many News as there are independent clauses in Fries' scheme, leaving no room for separate New information in dependent or embedded clauses. This differs from the approach in the current study that allows New information in dependent, independent and minor clauses and also analyses Theme in embedded clauses (see also section 3.1.2. on why the T-Unit may not be a suitable unit for discourse analysis).

When examining written English, Fries points out that there are a variety of potential systems available to realise the functions of information structure and New information. That is, although written English appears to have lost the options provided by the systems in spoken English such as intonation, stress and rhythm, it has found ways to compensate. Fries (1992, pp.462-463) outlines the following options for organising information in written English, so that focus can be realised:

- **Graphical**
  - physical placement in relation to the page
  - physical placement in relation to other items
  - size of lettering
  - style of lettering (font, colour, contrast)
- **Grammatical**
  - Clauses and clause complexes
- **Punctuation**
  - Boundaries of units
  - The marking of sentences
- **Sequential**
  - From known to unknown in a paragraph / text
  - Unmarked correlation between clause and information unit, and between end of clause and focus
  - From Theme to Rheme

It would be easy here to over-emphasise the **graphological** choices of underlining, italicising or bolding clauses, words or letters ("style of lettering" above) because they may be considered the main option in assigning contrast. These are, however, a highly marked option "used to mark only certain very important (usually contrastive) information" (Fries 1992, p.464). They are not an option that operates in every unit of information (roughly, clause), and may even be proscribed in some genres. While many of the options in this list have already been investigated, with particular emphasis being placed on grammatical and sequential aspects of meaning, it is punctuation that will be examined closely in the following sections. That is not to say that graphical options (or "multimodal" factors – see *e.g.* Kress and van Leeuwen, 1996; O'Toole, 1994; O'Halloran, 2004) should be ignored, only that they will not be reviewed in this study as they, again, represent highly marked forms of structuring information that cannot account for typical realisations in most texts.

Fries' more recent work has been invaluable in the development of the current study. Fries (2000; 2002) demonstrates clearly how the three systems of the textual metafunction operate independently, albeit with some unmarked correlations, to create meaningful choices. These papers divide the labour of the textual metafunction between the systems of participant Tracking, Theme and Information. In conclusion, Fries adopts a position that I have attempted to support throughout this study

> By contrast to the approaches of Chafe, Givón and Prince, I have argued for separating sharply the distinction between Given and New information in the information unit, as opposed to the distinction between presenting and presuming reference in the nominal group. (2000, p.103)

The current study has attempted through text analysis to demonstrate the independence of the systems of Reference and Theme, and their meaningful interaction with each other. The next section of the study attempts to demonstrate the independence of the third system - Information. Before that, however, we will investigate the results of SFL analyses of the interaction between Reference, Theme and information structure in academic text.

4.3.2 Information Structure and Academic Text

The analysis of information structure, and its relation to Theme and other grammatical systems, has been applied to academic text, both current and historical, particularly in Halliday and Martin (1993). Here we find refinements in the definitions of some major terms. For instance, Halliday defines New as "the element that constitutes the point of information for the message; this is signalled in English by nuclear prominence in the tone group" (Halliday and Martin 1993, p.60) and as "what you are to attend to." (p.90) Thus, although Halliday has made the function more explicitly distinct from reference, its realisation is still rooted in speech.

Halliday identifies the rhetorical functions of the unmarked combinations of Theme with Given and New with Rheme: "provided that the thematic element is also Given (i.e., non-New), the rhetorical effect is that of *backgrounding*" (Halliday and Martin, 1993, p.60, original emphasis). Similarly, the combination of New and Rheme functions rhetorically as *foregrounding*. The definitions for New and Given, however, depend on speaking the text aloud:

> Usually, the pattern of mapping of Theme + Rheme and Given + New on to one another is of this unmarked kind: the Theme is something that is given, and the New is something that is rhematic. This is especially true in written English, where (since there can be no tonic prominence until it is read aloud) the assumption is that the New matter will come in its unmarked position, namely at the end of the clause (Halliday and Martin, 1993, p.60).

By ultimately relying on reading aloud to define New information, the approach here does not account for the processes involved in reading silently (see 5.2). The realisation of New in written English is again not made particularly specific (in comparison, for example, to Matthiessen and Fries – see 4.3.1): New matter is described as being found at the end of the clause, but the extent of New is not made clear. This definition also suffers from the assertion that marked and unmarked meanings can be made, but is not clear on how

marked meanings would be realised; Halliday does not explain how New could be thematic, even when written text is read aloud.

Martin's approach, outlined in *English Text*, is applied in some detail to written scientific English in Halliday and Martin (1993). The difficulty in recognising the extent of New – the culmination of New can be identified in speech but, owing to its wave-like nature, the existence or extent of Given is indeterminate – is exacerbated in written English because "the tonic is not marked by English graphology" (Halliday and Martin, 1993, p.241). Martin's solution, as noted above, is to take a minimal approach "by taking New as the highest ranking clause constituent on whose final salient syllable the tonic falls" (Halliday and Martin, 1993, p.242). This definition ultimately requires the text to be read aloud, and so suffers the same problem of not being immediately applicable to most silent reading situations.

Halliday describes how Theme and New operate in conjunction with other systems in academic text to construe meanings. A crucial role is played by grammatical metaphor in packaging meanings in academic text. The congruent expression of a nominal participant, verbal process and logical conjunction are typically reshaped as nominal groups, thus enabling them to act as participants in a new process. This strategy is made more powerful in combination with the rhetorical effect of *backgrounding* nominalisations from the previous text and *foregrounding* New or focal nominalisations. When we look at this with the added perspective of Reference (see Table 4.3 for example), we find a powerful rhetorical pattern where the interpersonal Subject becomes largely a non-negotiable packaging of encapsulated meanings from the previous text and is followed by an incremental step towards a development in meaning. These developments are then encapsulated in a new nominalization, and the process repeats throughout the text. The result is a smooth *flow* of information in the text.

	The theoretical program of devising models of atomic nuclei	has of course been complemented by	experimental investigations.
Word class	(Encapsulating) Nominalisation	Process	(Developmental) Nominalisation
Theme	Theme	Rheme	
Reference	Presuming		Presenting
Information structure	Given		← New

Table 4.3 Flow of information in a clause (example from Halliday and Martin, 1993, p.92)

Lassen (2003) expands the work instigated by Halliday and Martin (1993) in academic text by synthesising SFL theory with Clark and Chafe to produce a description of how manipulation of Theme and information structure in technical writing excludes the novice reader, and in Lassen (2004) adds a critical perspective to analyse press releases from the biotechnology industry. Lassen (2004) describes how the deployment of grammatical metaphor makes implicit certain assumptions which, in their congruent form, would make explicit the agents responsible for various processes.

Martin (Halliday and Martin, 1993) identifies examples of Hyper- and Macro- Theme and New, and discusses how Themes function to predict, while News function to accumulate meanings in text (see section 4.2.1). Both of these resources are more easily identified, and often explicitly labelled, in written English than typically in spoken English – particularly conversation. This is a result of the nature of written text – it is not created in real-time, and can be reorganised in order to direct the reader. There are, then, certain aspects of written English that may be able to compensate for, or perhaps replace, features of spoken English that help the listener to find New information. In brief, analysis of academic text has revealed how clausal and discoursal systems of Theme and Information "conspire" with other features to rhetorical effect, but at this point New information can still only be identified when read aloud.

4.4 A Provisional Model of Information Structure in Written English

The discussion has reached a suitable point to propose a model for the SFL definition of Information Structure in written English, one which is realised in the clause alongside other systems in the textual metafunction. The following conclusions have been reached so far:

- Information structure operates independently of Reference.
- Information structure operates independently of Theme.
- Information structure is one form of prominence that operates alongside Reference and Theme.
- The linguistic functions and realisations described in most other theories as Information and Information Structure are accounted for in SFL by the systems of Reference and Theme, often in more detail.

- Information structure is not realised by intonation in written English because intonation is not realised in written English.
- Information structure must be able to function without resorting to reading aloud.
- The effects of information structure, as with Theme, will be influenced by the logical metafunction (see section 3.6).
- A model of information structure in written English should be able to identify the focus of New information at a level of delicacy approaching that of speech.
- Hyper-New and macro-New probably operate in parallel to hyper-Theme and macro-Theme.
- New information in written English is associated with final rhematic position – at the end of a clause or clause-complex.
- No description so far has described the possibilities for marked information structure patterns in written English; Theme and Information cannot be conflated as in spoken English.

A key issue here is the matter of the unit of analysis in written English. Davies (see section 4.3.1) maintains that punctuation offers only a rough guide to Information Structure, but this relates to spoken Information Structure which may indeed be different from written. Fries (1992, 1995a) argues for one information unit per T-unit – roughly an independent clause and all of its dependent and embedded clauses. Matthiessen's (1999) Culmination applies to each clause in a clause complex, as does Martin's (1992; Halliday and Martin, 1993) view of New (see sections 4.3.1. & 4.3.2). Combining Fries, Martin and Matthiessen, we can model information structure such that the final constituent in any written clause receives the focus of information.

It is likely that punctuation, as a marker of clauses, is likely to assist in dividing units of Information in written English. Halliday (1989) identifies three functions for punctuation marks, namely the marking of a boundary (typically a grammatical clause or clause complex), of status (*i.e.* speech function or speech act), or of a (logical) relation. Each of these options is likely to produce one distinct unit of information – regardless of the type of mark. We must then also include the possibility of clauses not being marked by punctuation, and decide whether this would count as a separate information unit. Since this study has allowed a Theme for each (dependent and independent) clause, we will assume that the same applies to Information. In written English, this allows common conjunctions to act in a similar way to punctuation marks in dividing clauses. That is, the

only marked option in the Theme-New system is when punctuation applies to non-clausal elements, creating units of information separate from a clause such as in a list where all items could be considered equal in informational value.

Bearing these points in mind, it would appear that the domain of New information in a clause is likely to be located in final position in the Rheme of each clause. More specifically, the more final a constituent, the higher its informational weighting is likely to be. That is, the pattern of increasing Information culminating at the end of a clause is a constant in written English. As this is still a working hypothesis, rather than name a linguistic feature New Information, we shall proceed by labelling them according to their realisation – in this case their sequence within Rheme. The analysis will distinguish clause constituents – *i.e.* groups, as in the other analyses in this study – that are final in a clause from those that are not thematic but not final, here labelled 'clause-internal', but will ignore groups that are indeterminate in their thematic or informational status (typically the verbal group in a declarative clause) if they are not in a position to compete for thematic or informational status. The resulting 'Final' units parallel Matthiessen's (1995a) Culmination and are similar to Fries' (1992) N-Rheme. The study will also investigate the distinction between what is final in a clause and what is final in a clause complex, producing the options in Fig. 4.4. That is, Information in a written clause is differentiated from Theme, and distinguished by relative position both within clauses and within clause complexes, resulting in a different realisation to Given and New in spoken English.

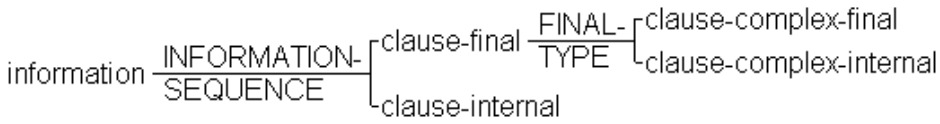


Fig. 4.4 System network for Information structure in written English

This study proposes that both Information Structure and Theme in a written English clause are realised by sequence. As Theme invariably starts a clause and New Information culminates at the end of a clause and clause complex, Theme and New appear to be mutually-exclusive in all cases of clauses with more than one word, preventing the two systems from producing marked examples where they are conflated. From another perspective, it would appear that Rheme and Given have no functional contribution to make to a written clause (see discussion in 3.5 on analysis of residues in the clause). Since both have no realisation rules, and are defined only as residues, Rheme and Given (in the informational sense rather than the referential sense – see section 4.1) can contribute little

to the analysis or generation of a written clause. At best, Rheme has been defined as the location or growth point for logogenesis, but the function of logogenesis in written English is largely carried out by New Information not by Rheme; Information flows naturally from New into Theme. Similarly, when Given is disassociated from its typical realisation by presuming reference, its indeterminacy contributes little to the analysis of a written clause, except to say that it is not what has already been defined as New information. As we have already introduced the concept that Information is gradable and increases throughout the written clause, Given means no more than 'less New than New Information'. This study accepts that Theme and Information waves typically work against each other in a clause, as in Fig. 4.5, and proposes that they do so as a rule in a clause in written English.

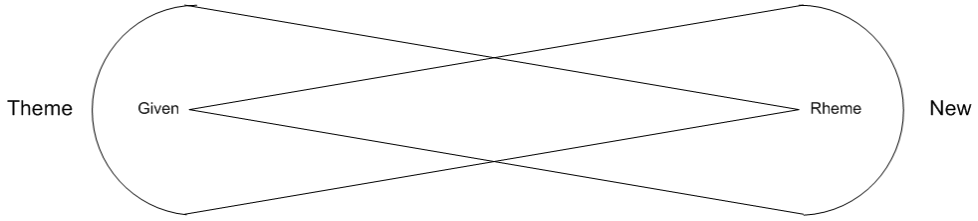


Fig. 4.5 **Opposing patterns of Theme and Information (after Matthiessen, 1992, p.42 and Halliday and Greaves, 2008, p.106)**

Consequently, Theme and Information Structure could be mapped onto the same system network for the analysis or generation of written English, as in Fig. 4.6. Despite claims of marked patterns, there have been no clear definitions of marked information structure in written English that do not depend either on reading text aloud or the function of participant identification (see section 4.3). Thus, a group in a written clause contributes to the Theme or Information Structure, or neither (typically in the 'pit' of the verbal group – see section 3.2). In generating a clause, a group may be assigned either a thematic or an informational position by the sequencing rules proposed by Hannay and Martínez Caro (2008). Both of these positions can be considered gradable, so that Initial Theme is more theme-like than Non-initial, and Final position is more New-like than Clause-internal, with an indeterminate area in the middle. Thus, part of the dynamics of written text is created by competitions to come first, for maximum thematic status, and to come last for maximum informational status. Clearly, this is a hypothetical claim that would need to be independently verified in a later study, although the model in Fig. 4.6 presumes this to be the case.

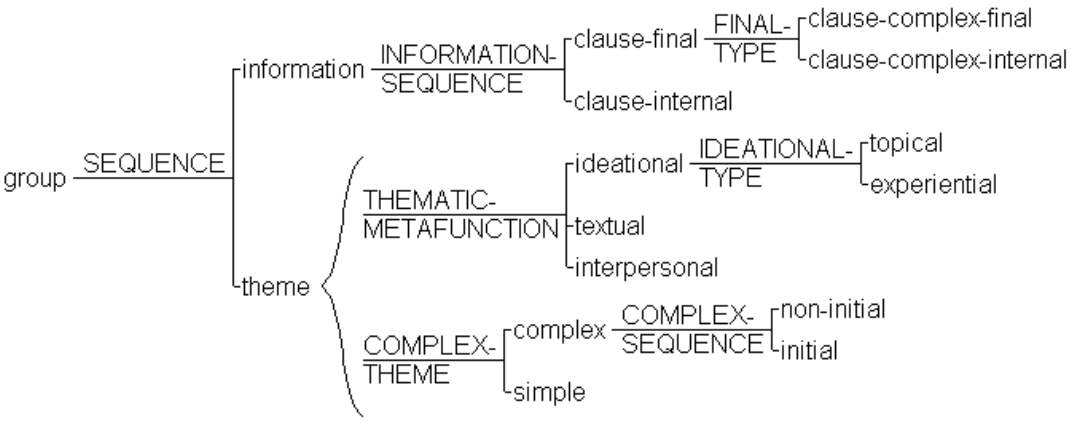


Fig. 4.6 Theme system with Clause-Final and Clause-Complex Final Information analysis

The scheme proposed in this study allows for a comprehensive model of the textual metafunction since it allows the three main systems of Reference, Theme and Information to interact to create meaning, albeit to a seemingly lesser extent than in speech. Even though the system in Fig. 4.6 appears to combine the two resources of Information and Theme, both of these systems can still interact with the systems of Participant Identification and Tracking. The implications of this change in textual dynamics for written English are discussed in section 4.7.

While clause-final position may often be taken up by a participant, realised as a nominal group, in this scheme it does not need to be. The scheme also reflects some aspects of Firbas' (1992) Communicative Dynamism (CD) in that it considers a range of factors – reference, sequence, and logical relations (see section 4.2.1.1). What makes this model different is that CD aims to identify one Theme and one Rheme proper in a clause or sentence. This model accepts that the three systems of Theme, Reference and Information all operate and interact, each making its own dynamic contribution to the different types of textual prominence in discourse. Section 4.6 provides examples. I would propose that when a single element realises Presenting reference and Clause-Complex Final New information, it would probably equate with Firbas' Rheme Proper (see section 4.2.1.1).

We also need to consider that, in hyper-New and macro-New, the thematically-driven macro- and hyper-Theme have their corollaries in the accumulation of News in clauses through a text (see section 4.2.1; Martin, 1992 p.456). This can be modelled in a system as in Fig. 4.7. In this system, only, the unit of analysis is the clause as a whole, rather than the groups in the clause. As with the analysis of hyper- and macro-Theme (see section 3.4.3),

clauses that function as hyper- and macro-New will also be identified. The wave-like organisation of the Textual metafunction predicts that written texts in particular genres will typically have patterns of New that encapsulate longer stretches of text (Martin, 1992; see section 4.2.1). Thus, a macro-New may encapsulate the hyper-New of a section of a text that may encapsulate the News in earlier clauses that may be adjacent or may be separated by local sequences of Themes and Information that are not related to the longer stretch of text. Information in a clause will be analysed according to whether meanings built up over previous clauses are encapsulated to create patterns of New, hyper-New and macro-New. It is hypothesised, however, that the texts are far more likely to exhibit characteristics of hyper- and macro-Theme than hyper- or macro-New due to the nature of the texts – each text is the introduction to a section of a textbook and so is more likely to prospect to meanings made later in the same section or chapter than to retrospect to meanings made earlier in the text. The analysis, as with groups in the clause, considers hyper- and macro-Theme and Information to be mutually-exclusive in written English.

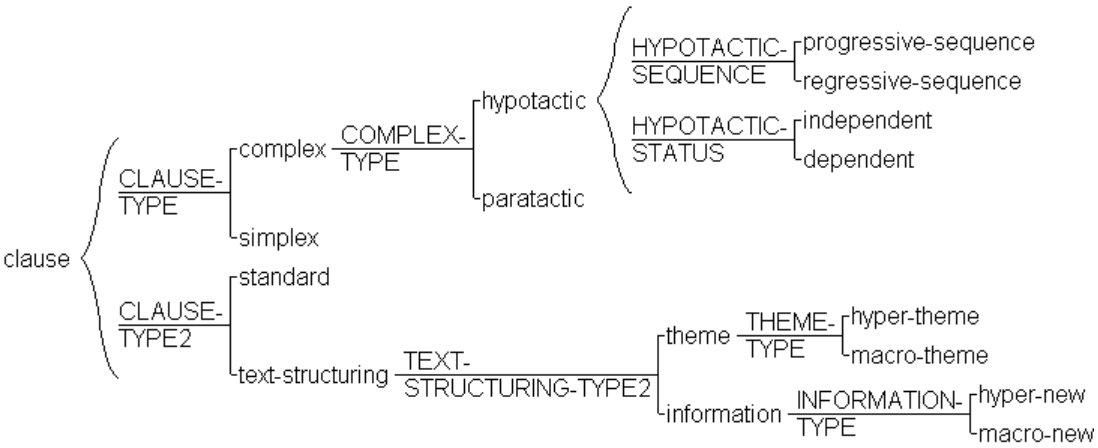


Fig. 4.7 System for analysis of Clause Logical Status and Role in Structuring Text

The system in Fig. 4.7 also considers the logical status of the clause, allowing for the identification of Progressive and Regressive sequences (see section 3.6). It thus proposes that there are different degrees of New information between a clause-final New in a dependent and an independent clause. Although the effects of the logical metafunction have been investigated in relation to Theme, I am aware of no current studies that transfer these effects to Information (or Rheme). However, there seems no reason to suggest that the observed effects on Theme would make no difference to Information. Taking into account the choices offered by the Logical metafunction, it is likely that within a clause complex, there is an incremental effect of sequence on the informational status of each

clause, so that a clause final group in a clause that is mid-complex will be less noticeable as New than the final New in the final clause of a clause complex. The differences between the effect of the logical metafunction on Theme and Information will need to be reviewed in further studies.

To investigate the interaction of the systems of Theme, Reference and Information, a model is required that can compare texts across genres, lengths and languages. A practical application which is presented here as a suitable model for analysis in this study is offered by Cummings (2005). Cummings reminds us that Rheme (*i.e.* Information) is equally as important as Theme when investigating discourse semantics. Consequently he builds into his quantitative model the interaction between Theme and Rheme. More specifically, Cummings (2005) tracks participants through Theme and Rheme to characterise the Method of Development of the text and develops a quantitative analysis based on reference chains (defined for this study as realised by Presuming reference). He compares the number of Themes that contain Presuming reference with the number of Fries' N-Rhemes (the final lexical group in an independent clause and all its dependent and minor clauses) that contain presuming reference, and counts the reference chains in the remainder of the clause, called 'Other'. Cummings' analysis lends support to the hypothesis that Theme and Rheme systematically select reference chains in different proportions, with Theme being more typically associated with presuming reference.

Typical results for a short expository extract are reproduced as Table 4.4 (Cummings, 2005 p.136). The first set of figures enumerates the proportion of all reference chains in Theme, N-Rheme or Other position (totalling 100%). The second set of results is the proportion of experiential elements in reference chains (or with presumed elements) in that part of the clause in relation to all experiential elements in that part of the clause. In this case, 77% of experiential elements in thematic position are part of a reference chain, while 60% of those in N-Rhemes are not. The final measurement is long chain distribution – the number of chains of reference that appear to be longer than others in the text. This arbitrarily relative measure probably needs a more consistent definition to be a reliable measure across texts. For this study, a long chain is defined as continuing for five or more repetitions. In Cummings' example, there is only one such long chain of 12 elements, of which 5 are in Theme in the 27 Themes that contain a reference chain – giving proportions, roughly, of 42% (5 elements in Theme of 12 in a long chain) and 19% (5 long chain elements in a total of 27 Themes that have chained elements). Multiplying the two provides an index that is comparable across different texts.

Chain element distribution in			Chain element density in			Long chain distribution		
Theme	Other	N-Rheme	Theme	Other	N-Rheme	% in Themes	% of Themes	Product
55%	20.5%	24.5%	77%	23%	40%	42%	19%	0.077

Table 4.4 "Table of proportions for the Elegant Universe text" (Figure 3 in Cummings, 2005 p.136)

The product of 0.077 for this expository text contrasts with Cummings' analysis of another short expository text with a product of 0.175, and even more so with extracts from the narratives of Dickens' "David Copperfield", with a product of 0.3, and Somerset-Maugham's "Of Human Bondage", with a product of 0.621. What Cummings appears to have achieved here is to quantify, and make comparable, the assumption that narratives make more use of pronominalisation and presuming reference than expository text, particularly in Theme position as a result of a small number of central participants that persist throughout the text. It also makes clear that within the categories of exposition and narrative there can be great variation. The figures for reference chains in an expository text (see Table 4.4), for instance, suggest that, with one exception, newly-introduced topics are not developed very far before another new topic is presented.

Both Martin (1992) and Cummings (2005) emphasise the complementary but distinct concepts of Theme and Information, producing the opposition of Method of Development and Point. The type of analysis used in this study for comparison between Theme and participant tracking, and for the identification of method of development will use Cummings' (2005) model, bearing in mind that some categories, such as N-Rheme, t-unit and 'Other' are only approximately equivalent, having been identified by separate criteria. In this study, the T-Unit will be replaced by clauses, so that every clause will be analysed for a Theme and for Information. The N-Rheme will be identified specifically, following Martin and Matthiessen (see above), as the last constituent in the clause, and labelled Clause-final. A further distinction will be made between the clause-final and clause-complex final positions and there will also be a distinction made between final position in a dependent or independent clause, allowing for comparison with Culmination and N-Rheme. The final rhematic group in a clause is classified as either Clause-final or Clause-complex final, the latter representing a group placed before the full stop of an orthographic sentence, whether placed in a complex or a simplex clause. Other, non-final,

elements in Rheme are labelled clause-internal (rather than 'other'), identifying potential candidates for final position.

4.5 Analysis of Information Structure

This section outlines the results of the analysis of Information structure based on the discussion in sections 4.1-4.4, initially using quantitative measures. Information structure displaces the Theme in a clause of written English (see Fig. 4.6), with New information culminating in final position. Information structure and New Information are analysed below before combining these analyses with the systems of Reference and Theme.

4.5.1 The Structure of Information: All Texts

There are more groups that contribute to the Information Structure of this corpus of texts than to Themes. Of the 2808 groups, 59.6% are in the Information portion of the clause. Non-participant groups are more likely to be found in Theme than in Information. With 60.9% of 647 non-participants in thematic positions in the clause, and 68.9% of 1809 participants in informational positions, this is a highly significant difference (χ^2 = (1) 177.6, $p<0.001$. See Appendix 4.1 Table 1). The higher number of informational elements in comparison with Theme is not trivial in the context of this study. With 867 Clause Internal groups competing with 807 groups to be in final position, the assumed focus of information for a clause, there appears to be a lot of pressure to 'come last'.

Information structure exhibits an almost identical distribution between groups in Final (807 groups or 48.2%) and Clause Internal (867 or 51.8%) position. This suggests an average of two informational groups (typically participants) in each clause – one clause-internal and one clause-final. When one of the 253 non-participants are realised in the information portion of the clause, there is a significant difference between appearing in Final position and participants in clause-internal position (χ^2=(1) 9.7, $p<0.01$. See Appendix 4.1 Table 2). 680 (54.6%) of the 1246 participant groups are found in non-final position. The 142 non-participants in Final position consist of 95 verbal forms (*e.g. occur*, *revolve*, *are*) of which 58 were participle forms, 61 adjectives[6] (*e.g. accurate*, *significant*, *highly efficient*) and 38 adverbs (*e.g. as follows*, *alone*, *further*). The figures show a discrepancy

[6] UAM Tools 2.0 (O'Donnell, 2009) offers a part-of-speech tagger (pos) based on a lexicon which assigns the most likely pos to an item. This may not always be accurate, and some words may be counted in more than one category (e.g. measured and lowered are counted as both participle and adjective).

between the total and the different classifications largely because some items were counted more than once –it was frequently difficult for UAMTools to distinguish between an adjective derived from a passive form or the passive form itself. For example, the UAMTools (O'Donnell, 2009) part-of-speech tagger recognises *measured* in the clause "When a quantity is measured" (EAP) as both adjective and participle.

Clause-Final position is divided fairly evenly between 408 groups (50.5%) that end a clause but not a sentence and 400 (49.5%) that end a clause complex (i.e. at the end of an orthographic sentence) – again a very close match between the two categories (see Fig. 4.8). Since every sentence must have a Clause-Complex Final element, it is interesting to note that in this genre there are almost the same number of Clause-Complex Internal groups, suggesting that, on average, each sentence contains two clauses. Typically, then, clause complexes in this genre average two clauses, each of which averages two groups in Rheme. There is no significant difference between the chance of a participant or non-participant coming in clause-final position within a clause-complex compared to the clause-final position at the end of a clause-complex.

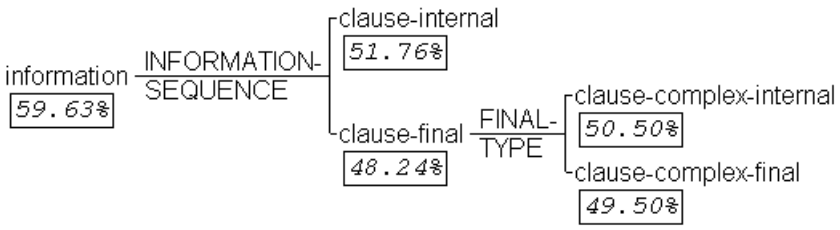


Fig. 4.8 Results for choices in all texts in Theme-Information network

4.5.2 The Structure of Information: Individual Texts

A comparison of information in clauses shows variation in typical complexity across the different texts in the corpus, as set out in Table 4.5. The groups in Final position in Rheme in column two are divided in columns 6 and 8 between clause-complex internal and clause-complex final. It must be remembered that clause complexes with more than one clause will contain both at least one clause final (which is not at the end of a sentence) and one clause-complex internal, and that clause-complex final also covers the final informational element in a simple clause.

The text with the simplest structure appears to be FSI (see table 1.2 for list of texts), with nearly 10% fewer clause-internal groups than clause final groups, suggesting that about 9

out of 10 clauses have just one group after the verbal group. Arc, on the other hand, has a ratio of two clause-internal groups to each clause final group, suggesting that clauses have an average of three groups after the verb.

	Rheme-Sequence				New-Type			
	final New		clause-internal		clause-complex-final		clause-complex-internal	
	n	%	n	%	n	%	n	%
Text 1: MM	81	47.1	91	52.9	45	55.6	36	44.4
Text 2: EAP	176	52.0	163	49.0	93	52.8	83	47.2
Text 3: BN	43	38.1	70	61.9	17	39.5	26	60.5
Text 4: Arc	33	32.3	69	67.7	12	41.4	17	58.6
Text 5: Sec	59	50.9	57	49.1	31	52.5	28	47.5
Text 6: AN	31	51.7	29	48.3	12	41.4	17	58.6
Text 7: RM	200	47.1	234	52.9	92	46.0	108	54.0
Text 8: FMAM	108	52.4	98	47.6	54	50.0	54	50.0
Text 9: FSI	79	59.4	54	40.6	40	50.6	39	49.4
Total	808		867		400		408	
Average		48.2		51.8		49.5		50.5

Table 4.5 The quantity and structure of informational Elements in different texts

Comparison of Clause-Complex Internal and Clause-Complex final results, which average a 50-50 split, produces results that can divide the texts into two groups: those with a relatively higher proportion of simple clauses (Texts 1, 2, 5, 8 and 9) and those with a relatively higher proportion of clause complexes.

The fact that information, unlike Theme, remains unanalysed for metafunctional role prevents a comparison with Theme for a Method of Development, or corresponding pattern, such as Culmination (Matthiessen, 1995a) or Point (Martin, 1992). Fries' (1981) notion of N-Rheme identifies the likely position for New information, but does not identify predictable patterns. No analysis has been attempted that identifies the semantic characteristics of New, or other, information. It seems feasible that an analysis that goes beyond the notion of complexity in Rheme into the Experiential content would reveal

similarly predictable patterns as those identified in Theme. This is an area that could be investigated at a later date.

4.5.3 Hyper- and Macro-New

Less obvious to identify were Hyper-New and Macro-New. This is partly a result of text selection - none of the texts were extracted from the end of their chapters or sections. In most cases, Information is specific and related to local co-text. Hyper-News are more likely to encapsulate earlier points, and are very likely to evaluate the text to that point. Compared to 106 text-structuring Themes (76:30 hyper:macro), there were only 15 text-structuring News (11:4 hyper:macro). In all cases, hyper- and macro-News required the reader to refer back to a previous part of the volume. For instance, in BN the sections starts with "Chapter 12 (section 12.3.2) introduced the concept of broadcast networks, and the most common topologies", encapsulating not the immediately-preceding section, but an earlier section being reintroduced to the discourse at this point.

4.5.4 Sample of the System of Information in a Text

A sample of one of the texts from the corpus will demonstrate how the system of New information operates from one clause to the next. For this exercise we use the text "The Nature of Data" (from EAP). In the following sample in Fig. 4.9, the New information for each clause is shown in a font which is one size larger, and clause complex final groups are shown in a font two sizes larger. This is done so that New information can be identified easily, and so that the reader can distinguish between a Final group that is Clause-Complex Internal from Clause-complex Final.

Taking only the Final positions, we find the sequence: *data – physical phenomena – ferric chloride – descriptive data – purely descriptive – qualitative data – measured – in a scientific way – present – quantitative data – the measurand – the human senses – the measurand.* Looking at Clause-complex News, we find: *data – physical phenomena – descriptive data – qualitative data – present – quantitative data – the measurand – the measurand.* In each case we see a development through the text, with the former list offering greater delicacy. There is also a lot of detail missing from the list of News, such as how each item may be related to each other. However, as an initial and very general overview of the text, the list of News offers some insight into the flow of information in the text.

2-1 THE NATURE OF DATA

Information you gather is called data. Data can be a factual statement of physical phenomena. For example, the statement "the copper was removed by the chemical reaction with ferric chloride"* is descriptive data. When data is purely descriptive, it is said to be qualitative data. When a quantity is measured, we associate numerical values with it, and the information is more useful in a scientific way because more information is present. Information about the magnitude or intensity of a physical phenomenon is called quantitative data. Recall that the quantity that is being measured is called the measurand. Instrumentation extends the human senses by allowing a numerical value or values to be associated with the measurand.

* This clause was analysed for Theme and Information because it is clearly projected, even though most embedded clauses were not analysed separately

Fig. 4.9 Information structure in sample text

Although the list of News for Fig. 4.9 may not give a full description of the text, we can discern some of its structure through the semantic relations between items. For instance, the first New, *data*, appears to be related, probably by a relationship of Hyponymy, to *descriptive data*, itself related to *qualitative data* probably in a Co-hyponomous relationship, which must then also be related by Complementarity to *quantitative data*. These relations, however, are part of the Participant Tracking system and are not strictly the concern of Information structure, although they help to explain how the semantic relations make the text appear coherent and while the presuming reference adds to its cohesion. The way that the New information is related textually, as we go through the text, is likely to be found in its interaction with the progression of Themes and the tracking of participants across clauses. Issues such as these will be discussed further in the following section when we look again at how this text operates with all three systems in a dynamic relationship.

4.6 Interaction of Information Structure, Theme and Participants

In this section we investigate the interrelationships between the three systems in the textual metafunction and examine the meaning-making resources that are available. The combinations available for meaning-making will be considered first quantitatively, before looking at the effects on a sample text a interpreting the results for implications in meaning-making.

4.6.1 Participants in Theme and Information in all Texts

This section reviews quantitative data on the interaction between the 3 systems being discussed. Results from a comparison of the Theme-Rheme system with Participant Identification revealed a significant correlation between Theme and Presuming in comparison with the correlation between Rheme and Presenting (section 3.9.1). In this section we will look at this relationship in more detail. Fig. 4.10 presents the proportions of choices in the system of Participant Identification, starting at Specified, for first and last positions in a clause: specifically, Simple Theme or Initial (in Complex) Theme (Set 1 in Fig. 4.10) or final 'New' position (Clause-Complex Final or Clause-Complex Internal) (Set 2). We have already discussed the role of participants and non-participants in Theme (sections 3.9.1 and 3.9.3) and Information (section 4.5.1). UAM CorpusTools 2.0 (O'Donnell 2009) reports no significant differences between Theme and Information for the initial choices of Effected-Neutralised or Specified-Generalized in the Identification network, largely because there are so few of the latter choices (0 and 11 of 1213 choices, respectively).

There are a total of 579 Participant groups in initial (i.e. Simple or Initial in Complex) thematic position and 634 in (clause or clause-complex) final position, or New Information, compared to 285 and 92 Non-participants, respectively, providing a highly significant association between Theme and Non-Participants and a negative association between Non-Participants and Information χ^2= (1) 90, $p<0.02$). (See Appendix 4.1 Table 3 for χ^2 contingency table.) Significant differences are also revealed between initial Theme and final New position in the systems of Presenting compared to Presuming (χ^2=(1) 25.93 or 25.34 with Yates correction, $p<0.001$), (Appendix 4.1 Table 4) and Presuming-nominal compared to Presuming-pronominal (χ^2=(1) 32.37 or 31.14 with Yates correction, $p<0.001$) (Appendix 4.1 Table 5). That is, pronominal groups, which are in the minority (99 compared to 557 nominal groups), appear more likely to be in initial (48% deviation

above the expected frequency) than in final position (57.8% deviation below the expected frequency), while Presuming-Variable-Nominal groups are significantly associated with final position (10.1% deviation above the expected frequency).

Significant differences to a level of $p<0.001$ also emerge comparing first and final position in the systems of Presuming-Undirected compared to Directed ($\chi^2=(1)$ 36) (see Appendix 4.1 Table 6). In this genre, contrary perhaps to expectation, there appear to be no significant differences within the Presenting system between Initial and Final position (Appendix 4.1 Table 7 gives figures for the Directed-Undirected choice).

Clear associations can be identified in the Participant tracking network, as can be seen in Table 4.6 which shows significant differences within the options for Participant Tracking between both Theme and Information, in general, and between initial (most thematic) and final (hypothesised New) positions in particular. For 1235 groups in Theme and 1444 in Information, there is a very highly significant association ($\chi^2=(1)$ 70.97, $p<0.001$) between a Non-Participant and Theme compared to the Information part of the clause (Appendix 4.1 Table 8). An Addition (no referent) is very highly significantly associated with Information compared to Theme, while a Referent has a very highly significant association with Theme (Appendix 4.1 Table 9). Specifying thematic position as Simple or Initial in Complex Theme, only, and Information as clause Final (Clause-Final or Clause-Complex Final) only, the same comparison of Addition to Referent gives $\chi^2=$ (1) 7.27 ($p<0.001$), suggesting a significant association between Non-Participants and the most thematic position.

Although the chi-square scores for Referent and Addition are similar, other types of Tracking relationship varied depending on how precisely Theme and Information were specified. While there were highly significant differences between Initial and Final position so that initial position (Simple Theme or Initial in Complex Theme) was strongly associated with Following reference ($\chi^2=(1)$ 24.51, $p<0.001$), the association between this phoric relation and Theme compared with Information was much less significant ($\chi^2=(1)$ 7.27, $p<0.01$) and appeared to follow the opposite pattern (compare Appendix 4 Tables 12 and 20 a and b).

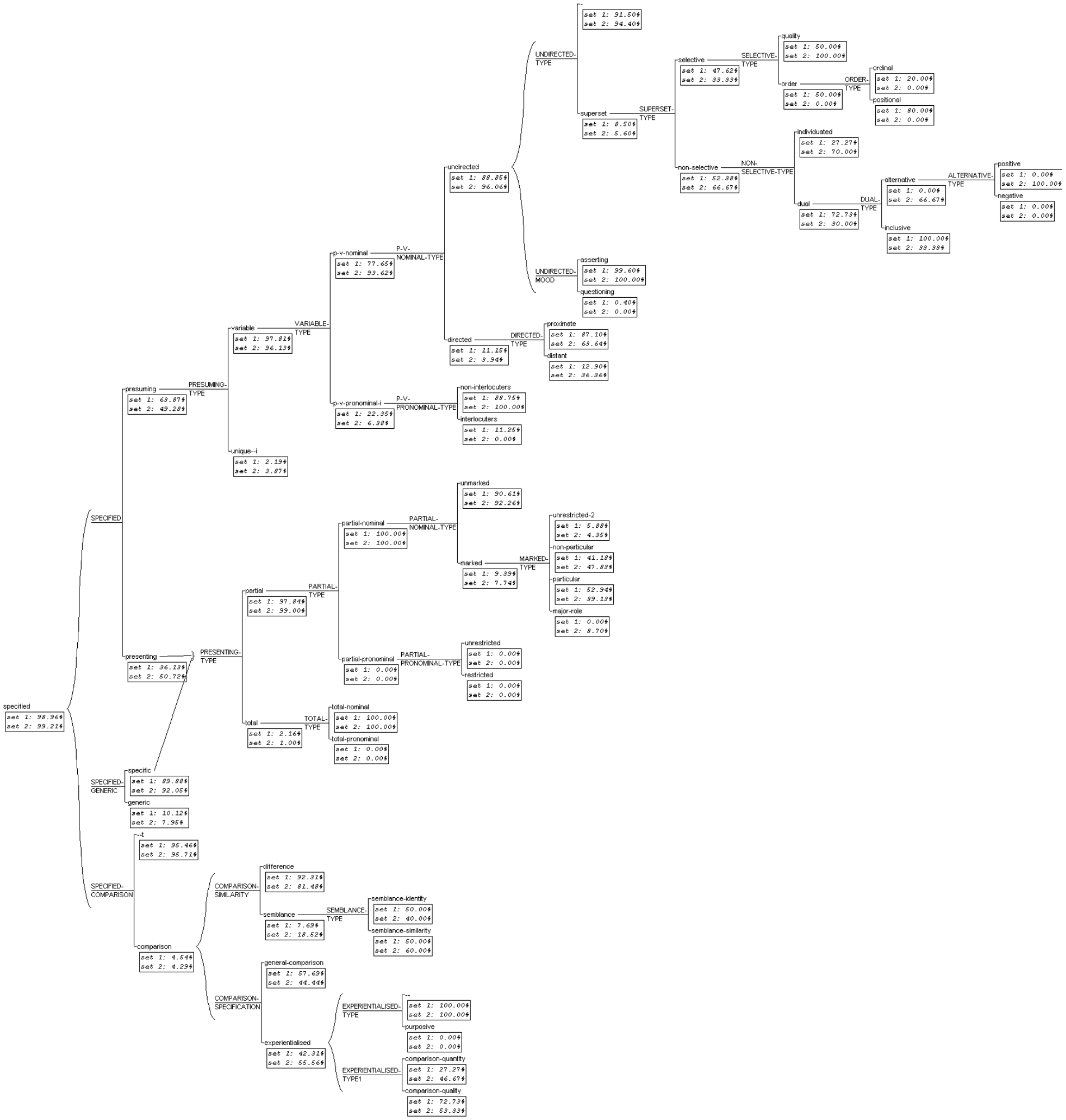


Fig. 4.10 Results of Initial-position-in-clause (Set 1) & Final-position-in-clause (Set 2) Proportions for Participant Identification Network

Participant Tracking	Theme *cf.* Information	Initial *cf.* Final position
Participant Non-participant	χ^2= (1) 70.97 (p<0.001) (Appendix 4.1 Table 8)	χ^2= (1) 82.96 (p<0.001) (Appendix 4.1 Table 16)
Referent Addition	χ^2= (1) 63.77 (p<0.001) (Appendix 4.1 Table 9)	χ^2= (1) 38.31 (p<0.001) (Appendix 4.1 Table 17)
Context of Culture (homophora) Context of Situation	χ^2= (1) 4.74 (p<0.05) (Appendix 4.1 Table 10)	χ^2= (1) 19.77 (p<0.001) (Appendix 4.1 Table 18)
Superordination Composition Nuclear Relations	χ^2= (2) 31.72 (p<0.001) (Appendix 4.1 Table 11)	χ^2= (2) 2.96* (Appendix 4.1 Table 19)
Co-hyponymy Class - Sub-class	χ^2= (1) 6.29 (p<0.02) (Appendix 4.1 Table 12)	χ^2= (1) 0.97* (Appendix 4.1 Table 20)
Complete Repetition Substitution	χ^2= (1) 40.76 (p<0.001) (Appendix 4.1 Table 13)	χ^2= (1) 33.68 (p<0.001) (Appendix 4.1 Table 21)
Preceding – Anaphora following	χ^2= (1) 7.27 (p<0.01) (Appendix 4.1 Table 14)	χ^2= (1) 24.51 (p<0.001) (Appendix 4.1 Table 22)
within group – esphora beyond group – cataphora	χ^2= (1) 0.098* † (Appendix 4.1 Table 15)	χ^2= (1) 14.43 (p<0.001)† (Appendix 4.1 Table 23)

* χ^2 showed no significant differences

† Results are unreliable due to low counts

Table 4.6 Summary of t-scores in comparison of Thematic and Informational status with Participant Tracking

Summarising these scores, the significant relations demonstrate that there are clear preferences in the selection of participants and non-participants and in the systems of participant Tracking and Identification that distinguish both Theme and Information in general and initial and final clause position in particular, at least for this corpus of texts. Similar analyses are required in this and other genres to establish the likely patterns of selection for these systems.

4.6.2 Participants in Theme and Information in Selected Texts

The interaction between Participants, Theme and Information Structure will be investigated here by closely analysing two texts from the corpus. MM (see table 1.2 for details) was chosen as it appears to be typical of the texts as it reveals average scores on most of the measures so far (see Appendix 3.3), and RM was chosen as the longest and the most atypical, at least in terms of register (Appendix 3.5).

In MM, out of 96 participants that are presented, 34 different participants are tracked for a total of 167 times. An analogy can be made here for Identification:Tracking (different participants:repeated participants) with the type:token ratio (different items:repeated items) in corpus linguistics, providing a ratio of 87:167 or 1:1.92. This figure can be compared with other texts to give an indication of cohesion. A small proportion (36%) of Presented participants were in Theme or in Titles (8%), while the majority of tracked participants were Presumed in Information (59%). That is, Presenting reference in Theme, at least in this text and probably for this genre, is a marked choice. In this typical example, a number of items are presented in the Information of the second sentence.

2.24 The knee, mounted on the column guideways, provides the vertical movement of the table. Power feed is available, through *a gearbox* mounted on the side, from *a separate built-in motor*, providing *a range of twelve feed* rates from 6 to250 mm/min. (Black, 1998, emphasis added)

Table 4.7 provides the distribution of participants in Theme, Information and titles. Table 4.7 and Fig. 4.11 both illustrate that, in this text, tracked participants appear in Theme and Information positions. For instance 11 participants were tracked once only, with eight of these repetitions occurring in Theme and three in Information, while one participant was tracked 18 times, of which 13 were in Information (see Fig. 4.11). Of the 34 participants tracked over 167 repetitions, the majority are tracked in Information. This may appear to cast doubt on the notion that Information is correlated with Presenting, and Theme with Presuming. The reason there is still a significant correlation between the two systems is that Rheme contains the majority of participants (61%) – both Presented and Presumed.

One interesting feature of this text is the tendency to introduce participants using a Title. In titles, the system of Participant Identification can be Neutralised, thus making no commitment on phoricity, which starts at the point of Presuming (see section 2.4.1.1). As titles do not do not exhibit the grammar of a clause, they can be interpreted as suspending Theme-Rheme structure. However, just as spoken English can realise Information

without clausal grammar, titles may still realise both New information and hyper-Theme, through sequence. MM often brings in a new participant without using Reference in a title. A subsequent Theme carries the same participant with presuming reference. That is, the Theme presumes the participant from the Title. This may cause difficulty to some readers and may also exercise linguistic power in assuming a context of culture through taking terms for granted before they have been explained in the text (Lassen, 2003).

Presumed Participants		Repeated in..		
Repetitions	Count	Theme	Rheme	Title
1	11	8	3	
2	7	6	8	
3	8	8	16	
4	2	3	5	
5	2	6	4	
6	2	4	2	
8	2	9	7	
11	1	4	7	
12	3	12	24	
13	1	4	9	
18	1	5	13	
Total Tracked	34	69	98	0
Total Tracked Participants				167
Presented Participants		Presented in….		
		Theme	Rheme	Title
		21	59	7
Total Presented Participants				87

Table 4.7 **Presented and Presumed Participants in Theme, Rheme and Title in the "Milling Machine" text**

In the second text that was studied in detail (a section of RM of comparable length), 53 participants were tracked 177 times through the text, making a ratio of 53:177 or 1:3.34. This suggests that Participants were tracked more often than in MM, partly as a result of not Presenting quite so many participants in a similar length of text, suggesting greater cohesion and more ideational consistency for the reader.

In RM, Participant tracking was very evenly distributed between Theme (49%) and Information (48%), while the 113 Presented participants were very strongly associated with Information (73%). This text used Titles to Present participants (11%) – almost as frequently as it used Theme (16%) for this purpose – but also employed Titles (3%) as a marked option to track participants. That is, although titles have no Theme-Rheme structure, they do not always suspend reference, and so participants can be presented in

title or text, and presumed in a later title. Table 4.8 provides a comparable table to 4.7 for this text, and Fig 4.12 complements Fig. 4.11. While most of the participants tracked up to five times reappear more often in Theme than Information, it is interesting to note (see Fig. 4.12) that participants repeated six times or more are more often presumed in the Information portion of later clauses than the Theme.

Presumed Participants		Repeated in …		
Repetitions	Count	Theme	Rheme	Title
1	22	12	8	2
2	8	9	6	1
3	7	8	12	1
4	8	23	9	
5	3	8	6	1
6	2	6	6	
10	1	3	7	
16	1	5	11	
33	1	12	20	1
Total Tracked	53	86	85	6
Total Tracked Participants				177
Presented Participants		Presented in…		
		Theme	Rheme	Title
		18	82	13
Total Presented Participants				113

Table 4.8 **Presented and Presumed Participants in Theme, Rheme and Title in the "Retardation Methods" text**

While both of these texts manage to weave a large number of participants successfully through the texts, RM appears to be more successful. Fig. 4.12 shows it has the higher count for the tracking of a single participant (33 repetitions compared to 18), and also repeats twice as many participants once, compared to MM (22 compared to 11). In the region of 2-10 repetitions the two texts are quite similar. The highest repetition in RM is in Information. The five most tracked participants are repeated 23 times in Theme and 44 times in Information, compared to the top 5 tracked participants in MM which are repeated 21 times in Theme and 56 times in Information. It seems that, over higher repetitions, RM is more likely to track participants through Theme than MM. This unmarked pattern is likely to be easier for readers to follow the ideas - even though they are expected to keep the same participant salient for longer.

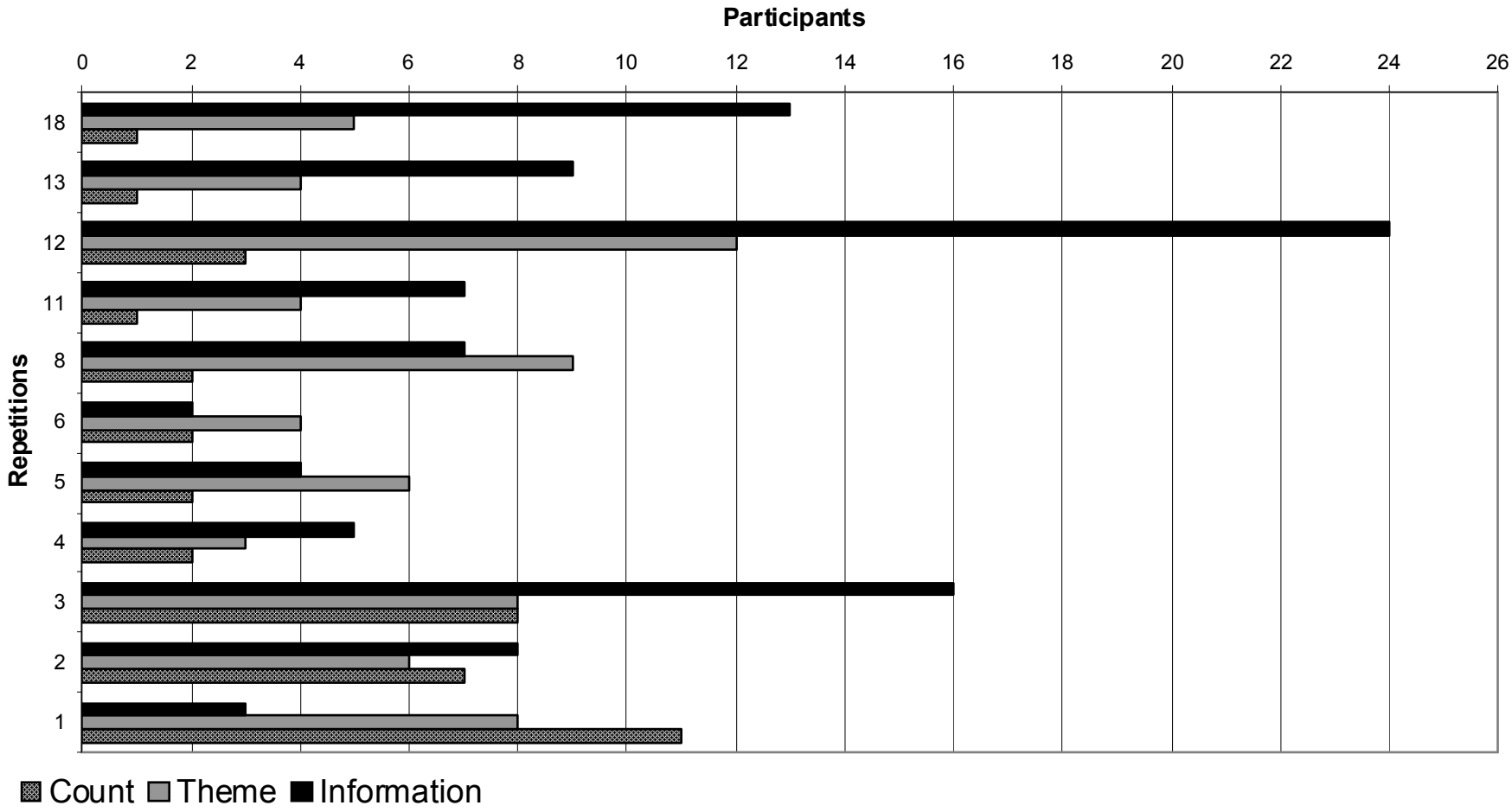


Fig. 4.11 Theme-Information Distribution of Tracked Participants in 'Milling Machine'

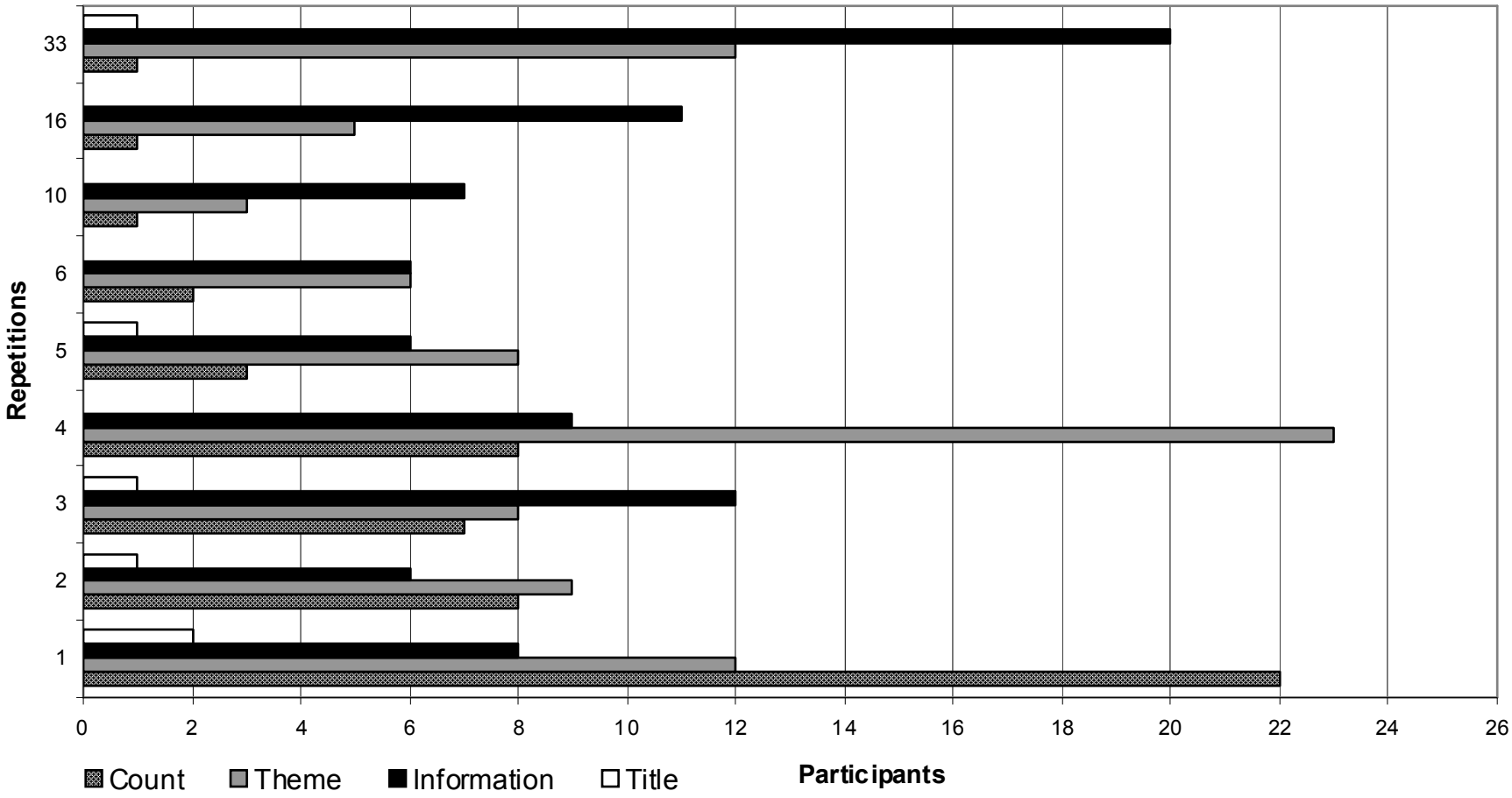


Fig. 4.12 Theme-Information Distribution of Tracked Participants in 'Retardation Methods'

4.6.3 Quantifying the Relationship between Theme, New and Reference

The analysis will now turn to the method proposed by Cummings (2005), outlined above in section 4.5. While attempting to follow the method as closely as possible, it must be remembered that the units of analysis are not exactly the same, and so the results will not be directly comparable, although it should be possible to suggest some areas of approximate comparison. It is hoped that further analyses will demonstrate that Martin's approach to discourse semantics produces more reliable results for measures of cohesion.

Cummings' (2005) quantitative measure (see section 4.4), intended to show the significance of participant tracking to Theme and Information, was applied to the MM and RM texts. While Cummings (2005) used the notion of reference chains, based on Halliday and Hasan (1976), this study has dealt with Presenting and Presuming reference and the system of Participant Tracking (Martin, 1992). This was preferred over Halliday and Hasan's original formulation largely because Martin's scheme is based on discourse semantics which is more likely to reveal insights about the discoursal nature of information structure than Halliday's clause grammar, even though information structure in writing is most likely to be realised within the clause. Table 4.9 displays the proportions for the MM text and Table 4.10 shows the same results for the RM text.

Chain element distribution in…				Chain element density in…				Long chain distribution		
Theme	Other	N-Rheme		Theme	Other	N-Rheme		% in Themes	% of Themes	Product
41%	27%	32%		77%	50%	73%		40%	60%	0.3712
		Clause	Clause complex			Clause	Clause Complex			
		15%	17%			34%	39%			

Table 4.9 Table of Proportions for "Milling Machine" Text (after Cummings, 2005)

Within the MM text, Presuming reference occurs 41% of the time within Theme, 32% at the end of the clause, and 27% in a place "Other" than these - or as it is called in this analysis "Clause Internal". Cummings (2005) uses Fries' (1981) concept of N-Rheme. To look more carefully at the structure of Rheme, I have made a distinction between the final element in a clause (Clause-Complex Internal) or a clause complex (Final). Looking at Table 4.9 we see that the score of 32% for N-Rheme is divided fairly evenly between Presuming Reference in Clause-final position (15%) and Clause-Complex-final (17%).

The measure for chain element density in this text – the proportion of Presumed to other participants – reveals scores of 77%, 50% and 73% for Theme, Clause Internal and Clause Final positions, respectively. That is, there is not a great deal of difference between the score for Theme and Clause Final. The table again reveals the difference between Clause-final and Clause-Complex-final with scores again evenly split at 34% and 39%, respectively. To measure Long Chain Distribution I have taken Cummings' (2005, p.138) example of five repetitions of the same participant as the cut-off point between "long" and "short" chains. In the MM text, there are 12 instances of participants with five or more repetitions. These 12 participants are repeated approximately 40% in Theme position. This represents about 60% of all Presuming Themes in the text. The product of these final two results is 0.371.

Similarities between the proportions for the MM and RM (see Table 4.10) texts seem to be confined to the first measure - that of Chain Element Distribution – with similar scores of 49%, 30% and 21% for Theme, Clause-Internal and Clause-Final, respectively. The measure for Clause-final and Clause-Complex-final is similarly evenly split, at 11% and 10% respectively. However, if we take the difference between Chain Element distribution in Theme (49% in RM and 41% in MM) and its distribution in Rheme (21% and 32%, respectively), we see that there is a far greater distinction between Theme and Rheme in the "Retardation Methods" text (28%), indicating a greater tendency to track participants through Theme than for the "Milling Machine" text (9%).

Chain element distribution in				Chain element density in				Long chain distribution		
Theme	Other	N-Rheme		Theme	Other	N-Rheme		% in Themes	% of Themes	Product
49%	30%	21%		56%	24%	41%		49%	40%	0.196
		Clause	Clause complex			Clause	Clause complex			
		11%	10%			20%	21%			

Table 4.10 Table of Proportions for "Retardation Methods" Text (after Cummings, 2005)

The next score, the proportion of Themes with tracked participants compared to all Themes, is 56% - quite a lot lower than the same score for the MM text. The same is true for Clause Internal, where 24% of this position in the RM text represents less than half of the 50% for the same position in the MM text. Final position also varies with 41% of Final positions in RM being taken by a Presumed participant. The split between Clause-Complex Internal and Clause-Complex-Final is again almost exactly even, at 20% and 21%, respectively. The analysis for long chain distribution, or multiple participant

tracking, reveals that of the participants that are tracked five or more times, 40% of the repetitions are in Theme. For all tracked participants, 49% are in Theme. The product of these two scores is 0.196, representing a far lower score than the previously-analysed text. The results for the proportions of these texts suggests that while the second text has a higher proportion of long chains in Theme, these do not represent as large a proportion of all the Themes as the MM text, which uses its proportionally fewer long chains more consistently for its Method of Development.

The quantitative measurement of Theme and Participant tracking benefited greatly from the method described in Cummings (2005). However, the texts analysed in this study are representative of the same genre, and Cummings' study revealed large differences between genres. It is necessary, therefore, to repeat the method described here, using Martin's participant tracking and Theme, with texts from other genres to attempt to repeat Cummings' findings. Cummings' approach to quantifying proportions in informational systems was chosen as a suitable point of comparison and the quantitative results recommend it as a suitable method of comparing texts.

4.6.4 Clause-by-Clause Analysis of all Three Systems in Interaction in a Text

We will now consider the same text as in section 4.5.6, but this time we will consider how all of the systems combine with each other clause-by-clause to make textual meanings. Fig. 4.13 presents the text, here marked only for clauses. There are 9 main clauses in 6 clause complexes, with clauses 4, 5, 6, 8 and 9 having one hypotactic clause in each (shown in standard SFL notation as α for the main clause and β for its dependent clause). Three clauses, 1, 3 and 8, contain single embedded clauses, marked here as x.1. Below is a commentary on the choices in the text.

In the first clause, there is a simple Theme (our starting point) of *Information you gather*. In the 'pit' between Theme and Information we have a simple verbal process of Relational: Identification (Halliday and Matthiessen, 2004 p.227), equating the first and last parts of the clause. There are no clause internal elements and so in the final-New position in the clause, (which is the Clause-complex Final because simple and complex clauses culminate in Clause-complex Final), we find *data* which is also Presented in the Participant Identification system (and Addition in Participant Tracking), giving it prominence in the Identification, Tracking and Information systems.

2-1 THE NATURE OF DATA

[1...] Information [1.1] you gather [...1] is called data. [2] Data can be a factual statement of physical phenomena. [3...] For example, the statement [3.1]"the copper was removed by the chemical reaction with ferric chloride" [...3] is descriptive data. [4ß] When data is purely descriptive, [4α] it is said to be qualitative data. [5ß] When a quantity is measured, [5α] we associate numerical values with it, [6α] and the information is more useful in a scientific way [6ß] because more information is present. [7] Information about the magnitude or intensity of a physical phenomenon is called quantitative data. [8α] Recall [8ß...] that the quantity [8ß.1] that is being measured [...8ß] is called the measurand. [9α] Instrumentation extends the human senses [9ß] by allowing a numerical value or values to be associated with the measurand.

Fig. 4.13 Clausal Structure of sample text (EAP)

A simple Theme, *Data*, starts the second clause, repeating the same lexical item as the final-New in the previous clause. However, the term is Presented (again an Addition in the Tracking system), showing how Reference is not determined by co-text but by speaker/writer choice, the writer here not requiring the reader to recognise *Data* as part of the context. The verbal group includes a modal expression but is, again, simply a Relational: Identifying process. The Final New is a complex nominal group, consisting of two nominal groups both of which are identified with Presenting reference and are Additions to the Participant Tracking network for this text. The question for the focus of information of New information is whether the Thing and Head of the nominal group are conflated and have been postmodified (*a GAME of two halves*) or whether the Thing is premodified by the Head (*a lot of TALK*). In this case *physical phenomena* appears to be premodified by *a factual statement* operating as classifier (see section 2.5 for discussion). Consequently, *physical phenomena*, as experiential Thing was judged to be New information in this clause. Although Halliday and Matthiessen (2004) suggest that "the unmarked focus of information in a nominal group is on the word that comes last, not the word that functions as Thing" (p.328-9), they also note that when Head acts in the group as measure or type expressions "it is often phonologically weak (non-salient)" (p.334) suggesting that in these and similar structures the Head will not carry the tonic foot and hence New

information. I would argue that readers can also recognise that in these expressions the Thing may be premodified by the Head, and so the New information in this clause is likely to be *physical phenomena* not because it is last but because it is the Experiential Thing.

The third clause starts with a Complex Theme. An initial Textual Theme (in unmarked position) *For example* is followed by the Topical Theme, itself containing an embedded clause with both Simple Theme *the copper*, Clause-internal Rheme *the chemical reaction with* and New *ferric chloride*. In this case, the embedded clause has been analysed in full, mainly because it is apparent from the text that this is a full projected clause. In this Complex Theme, all nominal groups carry Presuming reference. The Topical Theme *the statement* refers forward to the embedded clause which contains presumed *the copper*, and *the chemical reaction*. Neither of these items has been mentioned anywhere in the text, so their identity can be tracked to the context of culture (homophoric reference): all readers of this text are presumed to understand the general source of this quotation, and to see it as a 'citation' form. The term *ferric chloride* could be considered to carry Presenting reference. However, as it is clearly connected to the other items in the inverted commas it has been analysed here as Generic reference, again related to the context of culture. *Descriptive data* is the Clause (-complex) Final New information, and uses Presenting reference.

The first clause complex in our sample starts with a complex Theme in a Regressive sequence. The Initial Textual Theme *When* is followed by the Topical Theme *data* which still is not Presumed reference, (although it could be considered Generic, and so homophoric reference). The Clause-final *purely descriptive* is an example of a non-participant New, in this case in Clause-Complex Internal position. The main clause starts with the Simple Theme *it* which is Presumed with a Tracking relationship of Repetition-Substitution from *data* in the preceding clause. The New information for the clause and clause complex is *qualitative data*, which is Presented as an Addition to the tracking network despite the clear relationship of Contrast built up through the clause with descriptive data.

Clause 5 also starts with a complex clause in a Regressive sequence with the Initial Textual Theme *When*. The Topical Theme carries Presenting reference (*a quantity*). The New information in this clause is part of the verbal group, *measured*, resulting in marked Information for this corpus consisting of no nominal clause-internal elements and a non-participant New. The first main clause starts with the simple Theme *we*, a Presumed Interlocutor. In Clause-final New position is the Presumed *it*, which has a Tracking

relationship of Repetition-Substitution with *a quantity* in the preceding clause. The Clause-internal *numerical values* may have produced a better sequence, receiving clearer informational focus, had it been placed last: *When a quantity is measured, we associate it with numerical values*. A second main clause, in a paratactic relationship with the first and in a Progressive sequence with its hypotactic clause, starts with the Textual Theme *and* followed by the Topical Theme *the information*, which is identified as Presuming and can be tracked through Complete Repetition to *information you gather* in the first Theme. Again we find a non-participant in Rheme, this time the Clause-internal *more useful*. The New information in this clause is a Prepositional phrase containing a Presented participant *in a scientific way*. The final clause in the complex starts with the Textual Theme *because* followed by the Topical Theme *more information* – the first use of Comparison in the Identification network for this text. The Comparison is one of Difference and is Experientialised, and is also a comparison of Quantity. Because the repeated item has been experientially modified in this mention, it has been tracked though an Experientially Synonymous relationship. The Clause and Clause-Complex final New information is the non-participant *present*. Here is a prime example of the difference between the technical and the folk meanings of "Information". The word *present* adds very little in terms of new knowledge to the clause, but it is placed in clause-complex final position to focus on the importance of the culmination of the whole clause. An alternative may have been to use an existential clause, resulting in *because there is more information*, but this also culminates in a presumed item which in folk terms adds little new information.

Clause 7 starts with the Simple Theme *Information about the magnitude or intensity of a physical phenomenon*, although this is a very complex nominal group consisting of two Presented participants (*information* and *a physical phenomenon*) and two Presumed participants (*the magnitude* and (*the) intensity*) both of which can be tracked esphorically to *a physical phenomenon* in a relationship of Inalienable-Meronymy: Facet and Relational-Meronymy. There is no clause-internal group after the Relational: Identifying verbal group. The Clause (Complex) Final New information is the Presented *quantitative data*. Even though there is a clear lexical relationship of Dichotomy: Complementarity between this item and the earlier New *qualitative data*, the reader is not required to identify one with the other.

The Mood of clause 8 is marked for this genre – there appear to be very few Imperative clauses in the corpus. This results in a very marked analysis for this corpus. The Simple Topical Theme of *Recall* is in a hypotactic projecting relationship with the remainder of the clause complex. The lexical item *Recall* thus conflates Theme, Pit, Rheme and Clause-

final New, making it prominent in both the Theme and Information systems. The following, dependent clause starts with the Complex Textual Theme of *that* and Complex Topical Theme of *the quantity* which is esphorically tracked to *that is being measured*. (These Themes are then followed by an embedded clause which itself starts with the Textual Theme *that* and culminates in the Clause-final New *measured*. The analysis is presented here as a possible level of delicacy, but in fact the quantitative study left the embedded clause as an unanalysed part of the Topical Theme *the quantity* because the second layer of analysis contributes little.) After another verbal process of Relational: Identifying, the Clause-complex final New is *the measurand*. Although the item in New position is Presumed, the co-text makes clear that it is to be tracked either to the larger non-immediate verbal context or to the context of culture – as it has been analysed in this case.

The final clause in our sample is a hypotactic clause complex in a Progressive sequence starting with the Presented Simple Theme *Instrumentation*. Again we see the importance of writer choice in deciding both what to place as Theme and how to identify participants. The participant *Instrumentation* could easily have been introduced in Clause-internal or Clause-/Clause-Complex Final New. It could have been modified to include esphoric reference, such as *The instrumentation used for measurements*. The writer chose to leave the Theme as a Presented participant to be developed in the Rheme. In the main clause the Clause-final New is *the human senses*, Presumed from the context of culture. The following clause has a marked Theme in that there is no Experiential element, just the Textual Theme *by* which is sufficient to start the dependent clause. The process *allowing* contains a second part *to be associated* which is the third Clause-internal element. The other Clause-internal elements are the participants *a numerical value* and *or values*. These are Presented as Additions to the tracking network and left less in focus than the Clause-complex (and paragraph) Final New of *the measurand*. Looking at the choices available to the writer, clearly the elements in this final clause could have been changed to produce *by allowing the measurand to be associated with a numerical value or values*. However, the key term *measurand*, although it is seen as Complete-Repetition in the tracking network, receives the greatest focus here as the key term and as repeated from the previous clause.

4.6.5 Graphological Demonstration of all Three Systems in Interaction in a Text

We can illustrate the choices in the different systems, as described in the previous section, by using graphological devices to emphasise prominence in the texts. Fig. 4.14 uses the following scheme for the start of the FSI text: Themes are underlined, Presented participants are in **bold,** Presumed participants are in a lighter colour, all informational material is *italicised*, clause final groups are shown by a one size larger font, and clause complex final groups are shown in a font two sizes larger. Other texts can be found marked with the same scheme in appendices 4.2 to 4.4.

The analysed text has been presented in such a way as to suggest a recommended reading strategy. Unaltered text includes, in most clauses, the verbal process. These items will need focus only if the reader is lost or needs all details; academic text, in particular, often has very predictable verbal processes (Halliday and Martin, 1993). Bold participants need to be noticed more readily than faded participants. Underlined parts of the clause help the reader contextualise the most important focus of each clause, which is in italics. When New information is clause complex final and coincides with Presenting reference, it appears as bold italics in the largest font. This combination makes the item appear highly salient on the page, and is presumed to be an item that the writer wanted the reader to focus on.

In this text, then, the most salient items of New information are, in order: *a file?, an uninterrupted sequence of bits, different properties, a particular record, some field, quickly, a sequence of keyed records, access permissions, some of the attributes, created, (not) changed, immutable, changes.* Much as Method of Development charts an experiential flow of information through a text, so do these accumulating News, although this list of groups clearly can not tell the whole story of the text. Further studies are need, using this method, to identify typical patterns of culmination and accumulation.

Proceeding through the text, the graphological scheme reveals the interacting patterns of the flow of information produced by combining the Theme, Participant and Information systems. Using the Theme as a context, and the Information as the point of a clause we can see how the Participants and other elements are introduced and explored through the text.

5.1.1. The File Service Interface

For any file service, whether **for a single processor or for a distributed system**, the most fundamental issue is: *What is* ***a file*****? In many systems,** such as UNIX and MS-DOS, **a file** is ***an uninterpreted sequence of bytes***. The meaning and structure of the information in the files is entirely up to *the application programs;* the operating system is not *interested.*

On mainframes, however, **many types of files** *exist, each with* ***different properties.*** **A file** can be structured as ***a sequence of records,*** *for example,* ***with operating system calls to*** *read or write* ***a particular record.*** The record can usually be *specified* by giving *either its record number (i.e., position within the file)* ***or*** *the value of* ***some field.*** In the latter case, the operating system either maintains *the file as* **a B-tree** *or other suitable data structure,* or uses ***hash tables*** to locate *records quickly.* Since **most distributed systems** are intended for ***UNIX or MS-DOS environments*****, most file servers** support *the notion of* ***a file as a sequence of bytes*** *rather than as* ***a sequence of keyed records.***

A file can have ***attributes,*** which are ***pieces of information about*** *the file* but which are not ***part of*** *the file itself.* **Typical attributes** are *the owner,* ***size, creation date,*** ***and access permissions.*** The file service usually provides ***primitives*** to read and write ***some of the attributes.*** For example, it may be possible to change *the access permissions* ***but not*** *the size* (other than by appending ***data*** ***to*** *the file*). **In a few advanced systems,** it may be possible to create and manipulate ***user-defined attributes*** *in addition to the standard ones.*

Another important aspect of the file model is whether **files** can be *modified* after they have been *created.* Normally, they can *be*, but **in some distributed systems,** the only file operations are *CREATE and READ.* Once **a file** has been *created*, it cannot be *changed*. Such a file is said to be *immutable.* Having files be immutable makes *it much easier* to support ***file caching and replication*** because it eliminates ***all the problems*** associated with having to update ***all copies of a file*** whenever it *changes.*

Fig. 4.14 **Sample of text (FSI) marked for Participant Identification (Presenting / Presuming), Theme and Information**

In the first clause, the thematic context *For any file service* is completed with *a file?*. The Presented New is more salient than the Presented Theme, not least because it is asking a question, which appears to be answered in the second Clause-Complex. Within the thematic context of Presented *many systems*, the question in the first clause is answered as Presented New *an uninterpreted sequence of bytes*. The following clause appears to be low in participant and informational salience; all participants are Presumed, and *(not) interested* is the only New information, suggesting that this an additional and less important clause than previous clauses, with only the more specific thematic context changing.

The second paragraph provides the Presented thematic context of *On mainframes* and *many types of files*. The first New (Clause-complex Internal) is the non-participant process *exist*, which is then specified as *different properties* to match the files. The Presented Theme *A file* in the following clause is then given two Presented News – Clause-Complex Internal *a sequence of records* and *a particular record*. These participants then become the thematic context of the following clause as Presumed *The record* which receives its first New as the non-participant *specified*. A second thematic context *by* describes the relationship to the Clause-Complex Final New of *the value of some field*, which itself consists of Presented *some field* and esphoric *the value*. Arguably, as there is a hypotactic relationship between the last two items in this clause, the Presented *its record number* could also be considered as competing for final New position.

The following clause reveals further how the three textual systems described here operate independently. The Experiential Theme *In the latter case* and the Topical Theme *the operating system* are both Presumed, and the former is in Comparison with the previous clause. Both Themes are implied in the second clause through the elliptical use of *or*. The Clause-Complex Internal New *other suitable data structure* is Presumed and in Comparison to the clause internal Presented participant *a B-tree*. The next Clause-Complex Internal New is the Presented 'hash tables' and the Clause-Complex Final New is the non-participant *quickly*. To demonstrate how the systems could be reconfigured to give a different textuality, we can look at the same clause following a more 'canonical' pattern:

In the latter case, the operating system either locates *records* ***quickly using hash tables*** or maintains *the file as* ***a suitable data structure such as a B-tree***

In this case, the Presented *hash tables* and *b-tree* are placed in Clause-Complex Internal and Clause-Complex Final New information positions, respectively. Presumably, the writer of

the original used the dynamics of the three textual systems to provide a message that provides a greater degree of dynamic interaction between the three systems.

Without providing an exhaustive, and time-consuming, analysis of the whole of this text, I hope that this sample has demonstrated how the three textual systems of Theme, Participants and Information operate independently of each other in written English to provide meaningful choices for the writer.

4.7 Developments in Written Information Structure

Among the alternative descriptions to the SFL view of Information Structure, perhaps the most noticeable agreement is the desire to explain the choices behind deciding what to encode as one unit of information. Chafe, Steedman, Jackendoff, Prince and others, are all concerned to identify the psychological factors that make a speaker select a pattern of intonation, and divide a clause into a specific unit of information. The question is posed by Halliday: "The information unit is what the speaker chooses to encode as a unit of discourse" (1967a, p.202) but how the speaker chooses remains unanswered in SFL studies. This is an area that needs to be investigated carefully if the SFL perspective is to adequately account for the linguistic feature of information structure, in spoken or written language.

The following discussion proposes a development in the description of Information Structure as it has been presented and analysed in this chapter. Section 3.3 discussed "Special Thematic Structures". These include what are traditionally known as cleft and pseudo-cleft sentences. Although they have been referred to as thematic structures, it will be argued here that they in fact function to provide options for information structure.

Huang (1996) describes the functions of "enhanced Theme", as he names these structures, in terms of Prince's (1981) definition of Given and New (see section 3.3). That is, enhanced Theme is analysed through a model of reference (see section 4.1.3.1). However, in the Textual metafunction there is an important distinction between the Referential and Information systems (Fries, 2000; Martin, 1992). This distinction is blurred by Prince (from Clark, see discussion in section 2.2.1). Similarly, in Kaltenbock's (2005) review of *it*-clefts the definition of Given and New are associated with Retrievability, not in Halliday's sense of *presented as* Retrievable, but in the sense of Firbas'

FSP in that Given information is derivable from the con/co-text while New is not Retrievable within an "upper limit of nine preceding text-units" (Kaltenbock, 2005, p.127). Again, this study has revealed Reference to be independent of Information.

In an extensive treatment of clefts and pseudo-clefts, Collins (1991) describes the importance of these structures to information structure, particularly in written text. Collins dismisses the need for a psycholinguistic description of Given and New typical of Clark, Prince and their followers:

> it is not necessary to have recourse to the (somewhat inaccessible) notion of the 'hearer's consciousness' in order to explain the discourse behaviour of cleft constructions (Collins 1991, p.103),

a belief which he also applies to Gundel's (1985) work. His conclusion, however, depends on a referential distinction, as set out in Fig. 4.15 (figure 5.1 in Collins). That is, the 'degrees of informativity' are based roughly on Prince's taxonomy with the terms Inferrable and Fresh transferred directly. The main difference with Collins' scheme is that he relates each of the categories to the placement and intensity of tonic prominence, Fresh being the most salient by being the loudest, clearest change in pitch and Stale being realised by unstressed syllables carrying no change in intonation.

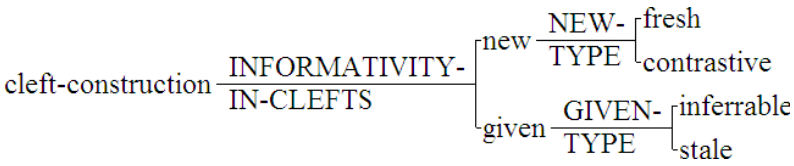


Fig. 4.15 System of Informativity in Clefts (Figure 5.1 in Collins 1991, p.110)

Collins' position that the realisation of information structure in speech is defined by Reference would be difficult to accept in this study. Further, he leaves the role of information structure in written English unanswered. He suggests that there may be a 'latent' pattern of information structure in written text realised by syntactic structure including the correlation between clause and tone unit, and punctuation. He alternatively suggests that "intonation plays a less significant role in the realisation of information structure in writing than it does in speech" requiring us to "accord a more central role to non-intonational phenomena" (1991, p.91) – precisely what this study has attempted to demonstrate.

Returning to Halliday and Matthiessen's (2004, p.97) example, we can see the two layers of analysis that are possible in these structures (see Fig 4.16). In this example, a

'congruent' sentence would read *His teacher persuaded him to come*. In spoken English it is quite feasible to produce this single clause with two information units, the first culminating in *teacher*, the second in *come*. In written English, however, this study proposes that the congruent sentence would be read as one information unit. Clefting a single clause in two produces two information units and two News in written English, as shown in layer i in Fig. 4.16. In the system provided in this study the Presumed participant *his teacher* would be Clause-Complex Internal New and the non-participant *come* would be Clause-Final New. The change in Theme is a result of producing two News. The ideationally empty, structurally cataphoric *it* in cleft structures is evidence of the lesser importance of Theme. Furthermore, in terms of logical structure, this analysis is now no different from a description of an independent clause with a hypotactic wh-relative clause in a progressive sequence (see section 3.6). The changes in Theme in these structures are a consequence of, not the reason for, the changes in information structure. It is, rather, the changes in information structure that provide the motivation for these structures.

	It	was his teacher	who	persuaded him to come
(a)	Theme	Rheme	Theme	Rheme
i	Given	New	Given	New
(b)	Theme		Rheme	
ii	Given		New	

Fig. 4.16 Thematic structures of clause with predicated Theme (adapted from Fig. 3-22 in Halliday and Matthiessen, 2004, p.97)

Consequently, I believe that the term Special Informational structures is more suitable than Special Thematic structures. It is possible that what will come **last** in a clause – what is New – is chosen **first** by the speaker (Hannay and Martínez Caro, 2008). Thus, Thematic position in an English clause is the consequence of what is chosen for final position. If this is true, similar consequences could be predicted to follow from the pressures placed on the written English clause.

If it is the case that Theme and Information are invariable and are consistently realised through sequence, then the dynamics of meaning-making in written English differ from those in spoken English. It is possible to vary Participant Identification and Tracking with Theme and Information in speech, but it is not possible in written English to produce the marked pattern of Theme with New and Rheme with Given (in the non-referential, informational sense). Theme – the starting point of the clause – contrasts with Information – what the writer wants the reader to focus on; the writer cannot ask the

reader to focus on the starting point (except by clefting a sentence and starting the clause with a 'dummy subject'). For example, a pronoun (typically Presumed) can occur in initial position – where it is the starting point of the message – or in clause-complex final position – where it is the focus of the message. In Final position the pronoun is New information. The pronoun may be referentially old, but the writer is forcing the reader's focus on it. Likewise the writer may force the reader's focus on an adverbial participle of a phrasal verb, a process or some other form of non-participant.

If Theme and Information share the same form of realisation, it would appear that there is less meaning-making potential in written English compared to spoken English which has separate realisations for the two systems. I would propose that many of the perceived differences between written and spoken English have developed to compensate for this apparent reduction in meaning-making resources. This section discussed the informational motivation for clefting a sentence, thereby adding a New to a clause complex. Other ways of manipulating clauses so that final position can be selected by the writer include the use of passive voice, the choice of process based on ergative roles and the rearrangement of adverbial elements in a clause. These are typical linguistic features that readily identify language as more written than spoken. Clearly, further investigation into the effects of New information on these features is required.

4.8 Implications

Within written English we can now demonstrate that three systems in the textual metafunction operate independently within the written clause, as suggested by Fries (2000). Reference functions to identify which participants are being presented to the context and which can be presumed and tracked through a text, and is mainly realised through determiners. Theme provides the starting point from which the clause-as-message can be interpreted. Finally, Information offers a method of directing a reader's attention to what is newsworthy. Both Theme and Information are realised by sequence, working against each other in clause and clause-like structures.

Information structure functions to direct a reader's attention to the ideational element that is the current expansion point of the text; to what is *News*. To put it another way, the discoursal goal(s) of the text – the point – is to be found at the end of clause and clause complexes. This position is significant because it appears to affect expectations of the

development of discourse across clauses, as described in the sample text. This is clearly a different function from the referentially-based definitions of information structure by Lambrecht, Clark, Chafe, Steedman and Firbas. I would recommend, as a result of the initial findings of the discourse analysis presented here, that future citations to Halliday (1967a) are clearly distinguished if they fail to use the term information structure to describe the key function of *Newsworthiness*.

4.9 Conclusion

An investigation into the different theories surrounding the notion and the term information structure has revealed a wide range of concepts and definitions. This chapter has defined a function for Information Structure in written English that is independent of Theme and Reference, just as it is in spoken English. The preferred definition for New information in written English is: what the writer focuses the reader's attention on. This appears to be realised in written English by final position in the clause and clause complex, but has implications for discourse beyond the clause. Analysis of the corpus of texts shows clear patterns of interaction between marked and unmarked correlations between the three systems described in this study (Reference, Theme and Information). We have identified a function and the realisation of information structure in written English that is independent of other systems in the textual metafunction, but still plays a role in assigning prominence within the clause. What we have so far failed to do, however, is explain why sequence has come to play the role of realising information structure in written English. This is the aim of the final chapter.

Chapter 5 Information Structure in Written English

5.0 Introduction

The previous chapters presented some evidence to support the view that there is a function of information structure in written English which is distinct from the functions and realisations of Participant Identification and Tracking, and Theme. Most SFL researchers consider information structure in written English to be realised by sequence, and we have gone some way to demonstrating how this may work in text. The previous chapters have attempted to provide further support for the metafunctional hypothesis in SFL by detailing how the three main Textual systems function and interact in written as well as spoken modes.

To summarise, there are three independent systems at work within the clause, each contributing to the textual meaning by adding various values. Participant Identification functions to separate the individual entities (participants), typically nominal groups, in discourse. Through the system of Participant Identification, participants are, typically through reference, given the value of Presented to or Presumed from the text. The system of Participant Tracking then details the location of the Presumed reference. Through phoric and lexical relations, a Participant can be tracked through text.

Theme is the systemic functional value that describes the contribution of the clause to discourse. Within each clause, the Theme is what the clause is about – logically, textually, interpersonally, ideationally or any combination thereof. Across clauses, the pattern of Themes contribute to the development of the discourse.

Within SFL, Information Structure functions to divide the stream of discourse into units, each unit representing one message consisting of an obligatory New and optional Given. New is the part of the message considered by the speaker or writer as Newsworthy – the part of the message that the listener or reader is asked to focus on. In written English, New information is realised by sequence – New information precedes a punctuation mark or the end of a clause.

The preceding three chapters have, I believe, demonstrated through discourse analysis the independence of function and realisation of the three textual systems operating in the

English clause, as well as outlining how they interact to create meaning within text. For the systems of Participant Identification and Tracking, and Theme, the realisation of features in written English reflects those in spoken English (although it is likely that the frequency of each will be proportionally distinct). Having established a possible model through a systemic functional analysis of text for the interaction between the textual systems within a clause, we can now look at the reasons why sequence in written English may have replaced the role of intonation in spoken English in realising information structure. It appears necessary to provide reasons why information structure has developed separate realisations in spoken and written English when the realisations for other systems in the textual metafunction are the same across the two modes. Evidence for the centrality of sequence to the role of information structure in written English has been demonstrated, but not explained, through discourse analysis. An explanation may be found through an examination of a variety of linguistic sources including functional, historical, social, neurological and psychological. The following discussion will attempt to use these sources to establish the reasons why sequence realises information structure in written English.

5.1 The Function of Information Structure

Before we go any further in our discussion, it is important to remind ourselves of exactly what we are looking for. The function of information structure is to divide the continuous stream of language into manageable 'chunks' or units of information. In speech, this is realised prosodically by an intonation contour. The item within each unit that is chosen by the speaker as *newsworthy* is given the most prominent position in the intonation contour to try to make the listener focus on that item; New information is realised by the tonic foot.

New information is one kind of value (Matthiessen 1992) offered by the Textual metafunction. Specifically, the textual value instantiated by New information is prominence. Other textual values include Referential, Thematic, Conjunctive and Action Sequence (Martin 1992). It will be assumed that all Textual values are available in both written and spoken modes, and that where realisation is equivalent in the modes, the form of realisation will be the same or similar. Where the form of realisation is not available to one of the modes, as in the case of information structure, it is assumed that the same function will be realised in another way.

As well as giving value to the elements produced in the Ideational and Interpersonal metafunctions, it is also the function of the Textual metafunction to simultaneously instantiate those elements. What we are looking for, then, is the function of information structure in English that is realised in the expression plane of orthography rather than in that of phonology. This function structures the stream of current, relevant discourse into units of information, and adds value to one part of the unit by making it prominent in a way that will direct a reader's attention toward it.

5.2 The Features and Functions of Written Language

This study investigates the role of information structure in written English from a Systemic Functional perspective. It is vital, then, that we look at written text from a functional perspective. This section details relevant functions of written text.

It is often pointed out that spoken language is both ontogenetically and phylogenetically prior to written language. Phylogenetically, written language develops as a consequence of social functions that are associated with settled rather than hunter-gathering societies (Halliday, 1989). The effects on society of written language, developed because of the need for language to be permanent for bureaucratic, financial and other record-keeping functions, have been documented by the anthropologist Goody, the linguist and social theorist Ong and the cultural theorist McLuhan, among others. Goody (*e.g.* 1986) and Ong (*e.g.* 1982) concentrate on the effects of literacy on society, while McLuhan (*e.g.* 1970) describes, not uncontroversially, the effects of writing technologies.

While it has been demonstrated that written texts can be employed both as tools of power and as objects of resistance, the development of literacy was generally commercial and utilitarian. In the European context, in conjunction with the development of written vernaculars (*i.e.* non-Latin) for abstract religious and secular texts, there was an increasing demand by merchants and businessmen for methods of record-keeping and communication (Briggs, 2000). Written language developed to mirror those aspects of spoken language that were in most demand and could be easily transcribed by contemporary technologies. Without these bureaucratic and commercial demands the much-lauded invention of the printing press (*pace.* McLuhan, 1970) may have passed unnoticed. It was the prevalent social conditions, which had already instigated the

development of new modes of thinking, reading, interpretation, debate and discussion, rather than the technology *per se* that ensured the rapid spread of printing technology (Saenger, 1997). Printing technology did not foster new forms of literacy, but responded to the prevalent demands. The functions of writing have evolved through a dialectical process into the forms that exist today. As much as these comments refer to literacy in general, they also refer to aspects of the written system. Each development created new conditions which enabled further changes. For instance, the development of punctuation in western European Latin and vernacular scripts dramatically shifted the responsibility for accurate interpretation of text from the reader to the writer (Briggs, 2000).

Spoken and written language develop in a dialectical relationship. Once written language 'takes hold', it may assume a higher status partly because access to the education necessary for literacy is often restricted to more powerful members of society. In some societies, the written form, once developed, may also become the preferred medium for sacred texts (Halliday, 1989). When written language is associated with higher social or religious status, many of the conventions of written registers may feed back into spoken registers, altering some spoken forms and their frequency. Written language expands the semiotic resources of a literate society, and so there is no reason why those resources cannot be exploited by spoken language. It is quite possible, for example, that the structures commonly associated with the manipulation of information structure in written English, such as 'wh-clefts', have become more widespread in spoken English as a result of their written functions.

Once written language has developed, it is likely to have an effect on the structures of spoken language. Saenger (1997) notes how the following linguistic features were changed in written French vernaculars:

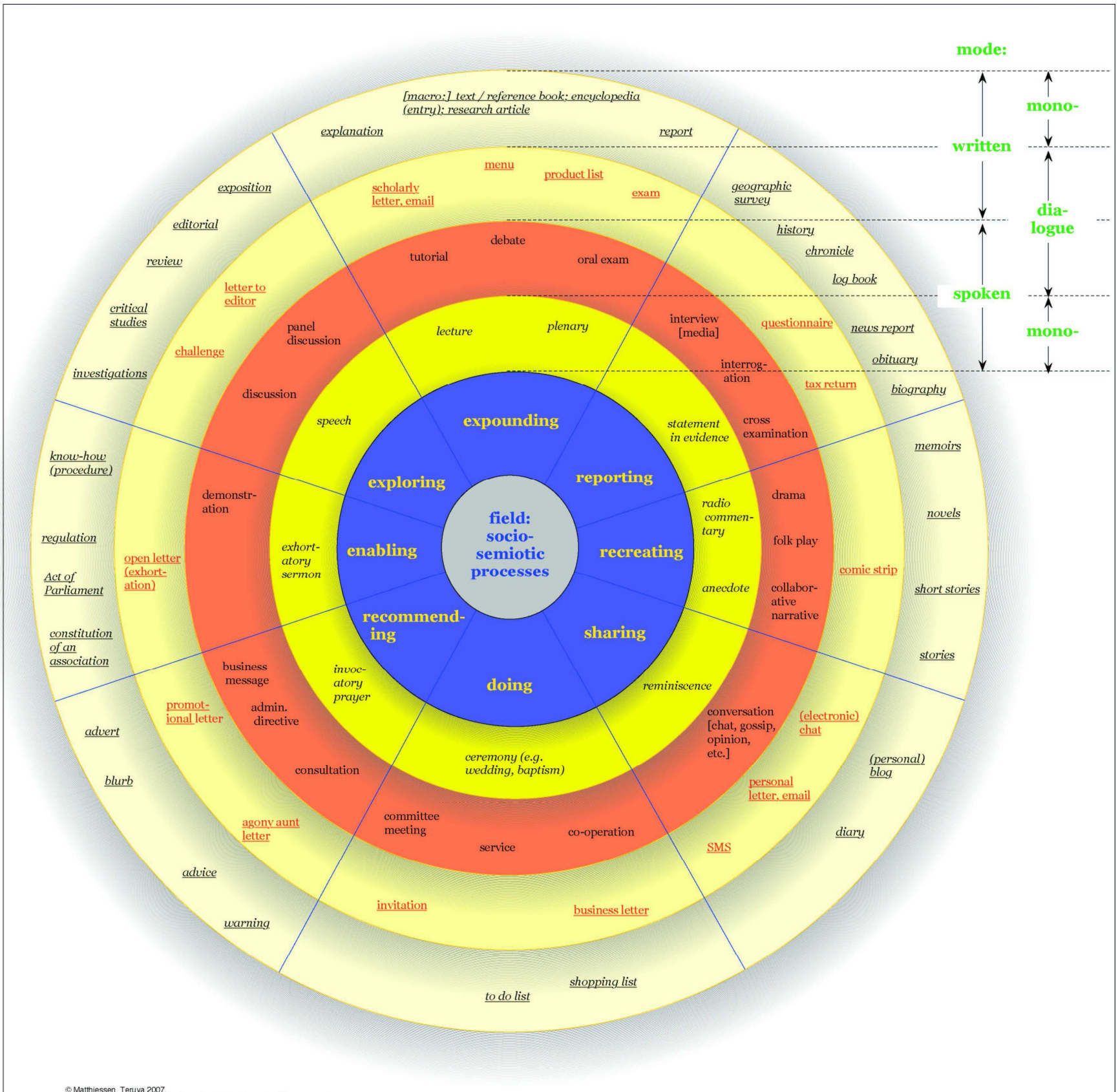
> Word separation, word order, emblematic punctuation, discrete clauses, the ordering of both words and clauses within complex sentences, and the use of conjunctions and adverbial conjunctions for the construction of compound and complex sentences (p.254)

Many of these changes produced changes in spoken vernaculars, particularly when the register demanded a style that reflected written language as this became more associated with a formal style. A special case concerns inversion for interrogative mood in French, which was invariably realised by intonation. While written French mirrored Latin and was written without spaces forcing its readers to speak aloud, there was no need for inversion to signal interrogative mood. With the introduction of spaces and the possibility of silent

reading, it became obligatory in written French to mark the interrogative with inversion. This has resulted in inversion also being used in spoken French (Saenger, 1997).

Matthiessen (2006) identifies distinct functions for written and spoken modes (see Fig. 5.1). The wheel of registers distinguishes in Field and Mode between spoken monologic, spoken dialogic, written dialogic and written monologic modes. Written and spoken text can also be analysed by an extra modal classification of intended mode of realisation. That is, some written registers are typically written to be spoken, while others are written to be read (just as some registers are spoken as if spoken and others spoken as if read) (Gregory and Carroll, 1978). For instance, most of the experiments carried out by Davies (*e.g.* 1986, 1994a) are of registers written-to-be-read, such as verse, newscasts and reconstructed commentary, and are spoken-dialogic in Matthiessen's scheme (Figure 5.1). Plays, speeches, poems and personal letters are also typically written to be spoken (or read to mirror speech). Spoken registers also vary in this respect with spontaneous conversation being the archetypal spoken-as-if-spoken register, with spoken situations that can be prepared, such as giving evidence or recounting a favourite anecdote or joke, representing examples of spoken-as-if-written registers, and reflecting features of written English such as greater macro-planning and lexical density.

The differences in writing and speaking result just not from the different modes, but from the different functions they are tasked to perform: "Writing and speaking are not just alternative ways of doing things; rather they are ways of doing different things." (Halliday, 1989, p.xv) Vachek (1987) describes the functional differences between the spoken and written norms of language thus: "the former [speaking] serves a ready and immediate reaction to extralingual reality, the latter enables the language user to react to it in a documentary and easily surveyable manner." (p.395) Written language must be functionally autonomous, but there must also be a correspondence between written and spoken language. Inevitably they will not be identical because of their different modes of realisation and rates of development, and their distinct functions.



© Matthiessen, Teruya 2007.
[This context-based text typology/ topology is based on Jean Ure's text typology; diagram by Christian Matthiessen & Kazuhiro Teruya.]

Fig. 5.1 Classification of Register by Field and Mode (by permission Matthiessen and Teruya, 2007)

The functions of written language require it to be more efficient than spoken language (Vachek, 1987). This increase in efficiency results in the separation of orthographic and phonological realisations of linguistic features. A highly complex system of notation would be required if written language were to recreate all of the features of spoken language used to identify individuals, including "rhythm, intonation, degrees of loudness, variation in voice quality ('tamber' [*sic.*]), pausing, and phrasing – as well as indexical features" (Halliday, 1989, p.30). It could be argued that this would dramatically reduce its efficiency in production and comprehension.

One can conjecture why, over the many centuries of the development of written English, no system of prosodic markings has been adopted. It would not be difficult, for example, to mark the focus of information with a preceding 'apostrophe, with underlining, or with some other gràphological feature (as can be seen in some folios of Shakespeare and as indicated here). If written language required the same realisation of information structure as spoken language it would be natural for a system to evolve. Since this has not happened, one assumption that could be made is that there is no information structure in written English. Chapter 4 established that this is highly unlikely.

Alternatively, it could be argued that there is in fact an inherent intonation contour in every written sentence. That is, information structure is realised by intonation even though no intonation contour is realised in written language. This contour can be identified whenever a sentence is read aloud, and so a tonic foot and New information can be identified for every clause. This is, indeed, the argument offered by, among others, Davies (*e.g.* 1989). Davies' subjects read aloud texts that generally mirror features of spoken English – verse, plays, newscasts etc. – and so it is concluded that readers are able to recover the original intended information structure of the text. Section 4.3.1 raised a number of objections to this view, including the use of text that is written-to-be-spoken, but this chapter will investigate whether this model is likely to work in all situations of reading – not just re-reading speech aloud.

In brief, features of written and spoken language develop in a dialectical relationship between the two modes as a result of the expanded semiotic resources provided by literacy. Phylogenetically, the functions realised by written language are distinct from spoken functions, and the modes develop different features in realisation, or at least in different proportions. While some written texts may be constructed to mirror features typical of spoken language, so that when read aloud the intonation patterns emulate

spontaneous speech, the lack of notation for intonation suggests that there is no necessary relationship between written language and intonation. This relationship will be investigated further in the next section.

5.2.1 Do we Hear Prosody when Reading Silently?

If we are to accept the argument that written English has an inherent information structure based on intonation, even though it is not realised until spoken, it would have to be assumed that when reading silently we are in fact reading 'aloud' in our head. On the other hand, we may propose that reading, particularly silent reading, is distinct from speaking. To posit a realisation of information structure in written text independent of intonation, we may assume that when we read we do not process language in the same way as when we listen.

The question we must ask, then, is whether we hear spoken language as we read. Davies describes the experience "that when we stop reading "the room seems to go quiet"" (1994a p.77) even in an empty room. Others problematise the issue: while phonological reading may be possible, "it is not *necessary* for word recognition and reading" according to Underwood and Holt (1979, p.82). From a functional perspective, Vachek (1989) insists that "No *detour* via the spoken utterances is absolutely necessary in deciphering" (p.100) written text. The debate makes it clear that it is possible to hear ourselves reading silently. That this is possible does not entail that it is necessary or consistent; neither does it indicate how frequently it may happen, nor how effective a reading strategy it may be. This section will review some of the research in this area, and make some provisional proposals that may help develop a theory of information structure.

For more than 30 years eye-movement research has improved our understanding of the physiological aspects of the reading process. Thanks to this methodology we now realise that what appears as the smooth process of reading is in fact the result of very fast movements by our eyes which appear to jump forward, rest and then move on again, with occasional back-tracking or "regression". This jumping is known as a *saccade*, while the time of resting is a *fixation*. Significantly, word boundaries – space between words – are crucial to a saccade. The spaces between words guide the eyes (Pollatsek and Rayner, 1982) and enable increased silent reading speed. However, not all words are equal and Rayner (1998) reviews a range of factors that influence whether a word receives fixation at

the end of a saccade. Chief among these factors is word frequency – both in general lexical frequency (as in corpus results for high-frequency words) and within the text, so that relatively frequent words are less likely to be a point of fixation. Rayner (1998) also lists semantic relations between words, morphemic units, anaphora and co-reference, lexical, phonological and syntactic ambiguity, and discourse factors and stylistic conventions as being studied for their effect on fixation. Other variables affecting fixation seemingly include word length, subject familiarity, concreteness and age of acquisition (Juhasz and Rayner, 2003). One hypothesis that would need testing for the purposes of this study would be the significance of the unit of the (verbal, nominal and other) group when reading silently. It seems possible that readers would focus on the Thing of each group, and on each instance of New information in a clause. Since none of the studies reviewed here operate with a SFL model of language it is not currently possible to verify this hypothesis.

Unfortunately, eye movement studies are inconclusive on the dependence on phonological versus orthographic clues when reading. Rayner, for instance, moves from claiming that phonological information is utilized early on in the reading process (Rayner *et al.*, 1998) to suggesting that phonological and orthographic clues support each other (Lee *et al.*, 1999) to a position that skilled readers do not need to utilize phonological information even during intensive 'conscious-level' tasks such as proofreading (Jared *et al.*, 1999). With an array of competing hypotheses using the same method of investigation, new methods would benefit us greatly. Eye-movement studies do not seem to show a great deal of reliability, probably because there is so much inference involved in transferring from observable eye movements to conjectured psychological processes.

Psychological studies that evaluate the role of phonological and graphological data when reading also produce conflicting conclusions. In 'priming' studies, Lukatela *et al.* (2001) concluded that reading depends on a sub-phonemic level of processing. They support the view that the processing of written words takes two simultaneous routes – one directly to a representation of lexical meanings, and the other via a phonemic path to the same destination. Lukatela *et al.* (2001) also suggest that "the visual word recognition system is even more intimately connected to the machinery of speech production and reception than heretofore recognized" (p.B48) and that separate phonological and visual systems may be working in tandem when reading words, as proposed by Coltheart and Coltheart (1997) in their review of cases where patients' linguistic ability was impaired. Perfetti and Bolger (2004) use functional neuroimaging to argue that visual, phonological and semantic

processes must be operating on written words. They caution, however, against any attempt at sequential or causational theorising based on neuroimaging results, and also demand more studies that look at language beyond the level of the word. There is, then, some evidence to suggest that when reading we may be able to hear words, and therefore that intonation might operate to realise information structure in both written and spoken English.

However, while possible, it may not be necessary to hear words when reading. Dehaene *et al.* (2005) review neuroimaging results and argue for a combinatorial system of reading that recognises combinations of letters as patterns, and learns the likelihood of those combinations. However, Dehaene *et al.* (2005) are referring mainly to the ability of a reader to recognise pathological words as real words; subjects recognise '*grdn*' but not '*gdrn*' as 'garden'. Brysbaert and Praet (1992) conclude that there is evidence "for the dominance of the orthographic route over the phonological route in processing isolated words." (p.91). It could be argued that if orthography is more significant than phonology for isolated words, it is likely to be even more so for connected words, sentences and texts. Combining a range of methodologies and analyses, Peereman *et al.* (1998) conclude that "the experimental observations reported in this paper cast doubts on the existence of reciprocal constraints between orthography and phonology at prelexical stages of processing." (p.171). That is, they see little evidence for the necessity of a phonological level of processing in reading.

While the studies mentioned above may support the view that phonological and orthographic systems can operate independently, they all have one major flaw in common. They all study isolated words, and presume that citation forms of words in experiments reflect meaningful use of words in context, an assumption that demands verification (Cutting *et al.*, 2006). That is, they lend some indirect support to the view that it is not necessary to sound words when reading silently, but more importantly demand research that investigates units larger than the word.

Looking at reading above the level of the word, Caplan (2004) reviews a range of neuroimaging studies related to sentence processing. One thing that is made apparent in this review is the narrow approach taken to language and comprehension in all such studies. In terms of SFL, very few studies venture beyond Ideational grammar; the majority of studies focus on predicate-based syntax. Despite clear intentions by some experimenters to prove the existence of an autonomous syntax, Caplan's review concludes

that not only is there no identifiable difference in the brain between lexical and syntactic processing, but there is no identifiable difference between hearing and reading. That is, although differences between word recognition tasks may be identified in neuroimaging studies, the same cannot be said for sentence comprehension tasks.

Starting with the visual system itself, Magnuson *et al.* (1998) increased the complexity of the chunking demanded of their subjects, resulting in visually highly-complex stimuli. Unfortunately, this study dealt mainly with abstract shapes and relations rather than letters and words. However, it seems probable that if the visual system is able to chunk large amounts of information into a single unit, then the process of reading will be able to draw on that ability. Conklin and Schmitt (2008) provide support for the importance of chunking information in their study of sequences of formulaic language. Looking specifically at reading, they conclude that formulaic sequences enable faster processing for both native-language and foreign-language readers.

A significant criticism of these studies that must be considered is the influence of the tools over our way of thinking. Much of the evidence we have for the phonological and non-phonological involvement in reading derives from neuroimaging studies using fMRI (functional magnetic resonance imaging) and PET (positron emission tomography) scans of the brain. The way these tools work appears to highlight one area of the brain in detail in order to measure the effect of particular stimuli. While there may be some value in this approach, and some interesting observations have been made, it is vital that we remember that these tools are highly localised – they view, in detail, what happens in one particular region of the brain at a very exact time. It is unsurprising, therefore, that results from these studies support the view that reading processes are executed locally and sequentially in the brain. As Edelman (2004) points out, however, a view of the brain that posits localised functions without considering the complex interconnected structure of the brain is now "indefensible" (p.30)

Salmelin and Kujala (2006) use different tools to arrive at a conclusion that they warn must only be seen as provisional. Developing the concept of a neural network, and exploiting MEG (magnetoencephalography) results, Salmelin and Kujala were able to trace the possible pathways that language, in many cases written language, takes through the brain. Among the most interesting findings are the revelations that there are a wide variety of areas and processes associated with other activities involved in the process of

reading, and that areas of the brain associated with both general visual processes and language production are stimulated when reading.

One interpretation of these findings would appear to support the view that we hear when we read. This could be considered a rather one-dimensional description, however, based on a view of the brain as operating with one location being associated with one function. It is possible to identify different functions for different parts of the brain, but this may be a result of methodology, telling us more about the investigator and mode of investigation than the phenomenon itself; tools which locate brain activity produce descriptions based on location. However, if we move from this cognitivist model to a more embodied, phenomenological perspective, as proposed in chapter 1, we may be able to incorporate these results into a more satisfactory framework.

It is possible that reading, as a linguistic process, does indeed activate those areas of the brain that we would use to motivate articulation and those areas that are associated with aural perception of language. That is to say, while some areas of the brain are particularly active for certain linguistic processes, the other areas do not shut down – there is a simulation, echoing or mirroring effect in other parts of the brain (Barsalou, 2008). That is, the brain operates amodally, in contrast to cognitivist theories that support a modal model of the brain required for compartmentalised systems of syntax, semantics and pragmatics. The brain 're-enacts' processes experienced through the learning of a behaviour or concept (with concepts being treated as internal simulations) (Barsalou, 2008). Thus, learning (including language learning) takes place through associations with a wide range of bodily movements and processes, and when that learned behaviour is called upon again, all of those associations are also recalled. When recalling a particular phrase, for example, the articulatory process required to pronounce the phrase is simulated in the brain even if the physical articulation is not enacted.

An amodal perspective, supported by TMS (transcranial magnetic simulation) results, reveals that when reading *action* words, such as 'kick', or related nouns, such as 'foot', the areas of the brain related to these motor functions were activated (Pulvermüller *et al.*, 2005). This effect has been described as an action-perception loop (Vigneau *et al.*, 2006) and is carried out entirely within the brain, with brain activity reflecting that of a physical process, even though the physical aspects of the system are not activated:

> The mental simulation of a motor act that is not accompanied by an overt body movement … corresponds to a process by which the brain activates

> a motor plan and monitors its unfolding through internal feed forward models, while holding back (overt) motoneuronal output. (Tomasino *et al.* 2007 p.T128)

It is this effect that probably produces the experience of 'hearing' what we are reading as much as, or more than, a conscious effort to sound out the written language in our head. In fact this effect is so great that Tomasino *et al.* (2007) contend that the type of language that is 'experienced' (action words, physical objects, abstractions etc.) has a more significant effect on brain activity than the mode, or channel, through which it is experienced.

According to research reviewed here, it may be possible to read without hearing the words in our head, and it may be that the effect is mediated by factors such as task or time. On the other hand, it may be that we cannot stop our brains from simulating the processes in the brain related to listening and speaking when reading. However, even if the brain simulates listening or speaking, the main point is that the constraints of the articulatory system are still negligible because there is no physical reaction – any motor-neuronal activity remains in the brain. The action-perception loop takes place within 'cranial time' at the speed of neuro-chemical transfer, and is thus unconstrained by the limitations of physical processes. Consequently, whether or not we 'hear' when reading, we can probably ignore the physical constraints of articulation and exploit the potential, or affordances, of the visual system. That is, the units of analysis, or Information units, in speech and in reading do not need to be the same, as they have the physical and cognitive constraints, and potential affordances, of different sensory systems.

5.2.2 The History of Silent Reading

If silent reading is not bound by the real-time constraints of articulation, but depends on eye movement (Pollatsek and Rayner, 1982) (see 5.2.1) for realisation, we need to look in detail at reading, and in particular the practice of silent reading – the normal mode of reading today – in order to accurately identify the differences between speaking and writing and to identify the need to voice and produce intonation when reading. Written English that is intended to be read silently is likely to demonstrate an information structure that need not concern itself with physical constraints such as drawing breath, resulting in relatively short regular 'pulses' of prominence. Saenger (1982) notes that as punctuation developed it "was calculated to guide the eye rather than to regulate the voice of a professional reader." (p.409) It is from the visual-oriented approach that those

written genres that are intended to be read silently have evolved such features as complex noun phrases – features that cannot be spoken comfortably in a single breath, but which cause the reader little or no difficulty because of the grammatical markers provided to visually pattern text into suitable chunks of meaning.

Evidence for the centrality of sequencing, chunking and punctuation to silently reading written English comes from exhaustive study by Saenger (1982; 1997). For more than 10,000 years, from ancient Greece and Rome to the courts and monasteries of 10th Century France, the practice of *scriptura continua* was the norm. That is, written language was a continuous stream of letters, as illustrated in Fig. 5.2 and exemplified by Halliday's (1989) modern version of the original Greek practice of writing uninterrupted letters on one line left-to-right followed by one line right-to-left:

> onethingwascertainthatthewhitekittenhadhadnothingtodowithititwa
> cafstignivahneebdahnettiketihwehtrofyleritnetluafsnettikkcalbehts
> ewashedbytheoldcatforthelastquarterofanhourandbearingitprettyw
> feihcsimehtnidnahynadahevahtndluoctitahteesuoyosgniredisnoclle
> (p.33)

While it may be possible, with practice, to 'parse' this stream of letters into text, the process is far simpler when reading aloud. It is the spaces that enable eyes to quickly identify potential units in peripheral vision and then saccade to the next focal unit:

> Without spaces to use for guideposts, the ancient reader needed more than twice the normal quantity of fixations and saccades per line of printed text. The reader of unseparated text also required a quantity of ocular regressions for which there is no parallel under modern reading conditions (Saenger, 1997 p.7)

Spaces were introduced between words into Latin texts in Ireland partly because of difficulties experienced by the non-romance speaking Irish monks in identifying transcription units in *scriptura continua* (Saenger, 1982; 1997). *Scriptura continua* (which by now followed the left-to-right convention for all lines) forced the reader to speak the lines in order to make sense of them:

> The Roman reader, reading aloud or softly to himself, approached the text syllable by syllable in order to recover the words and sentences conveying the meaning of the text. (Saenger, 1982, p.371)

That is to say, information units and information structure could only be recovered through reading aloud. Texts that were written to be read aloud produced their own information structure that closely followed the patterns of spoken English as the same

limitations of the lungs and speech organs operated on these texts as operated on real-time speech.

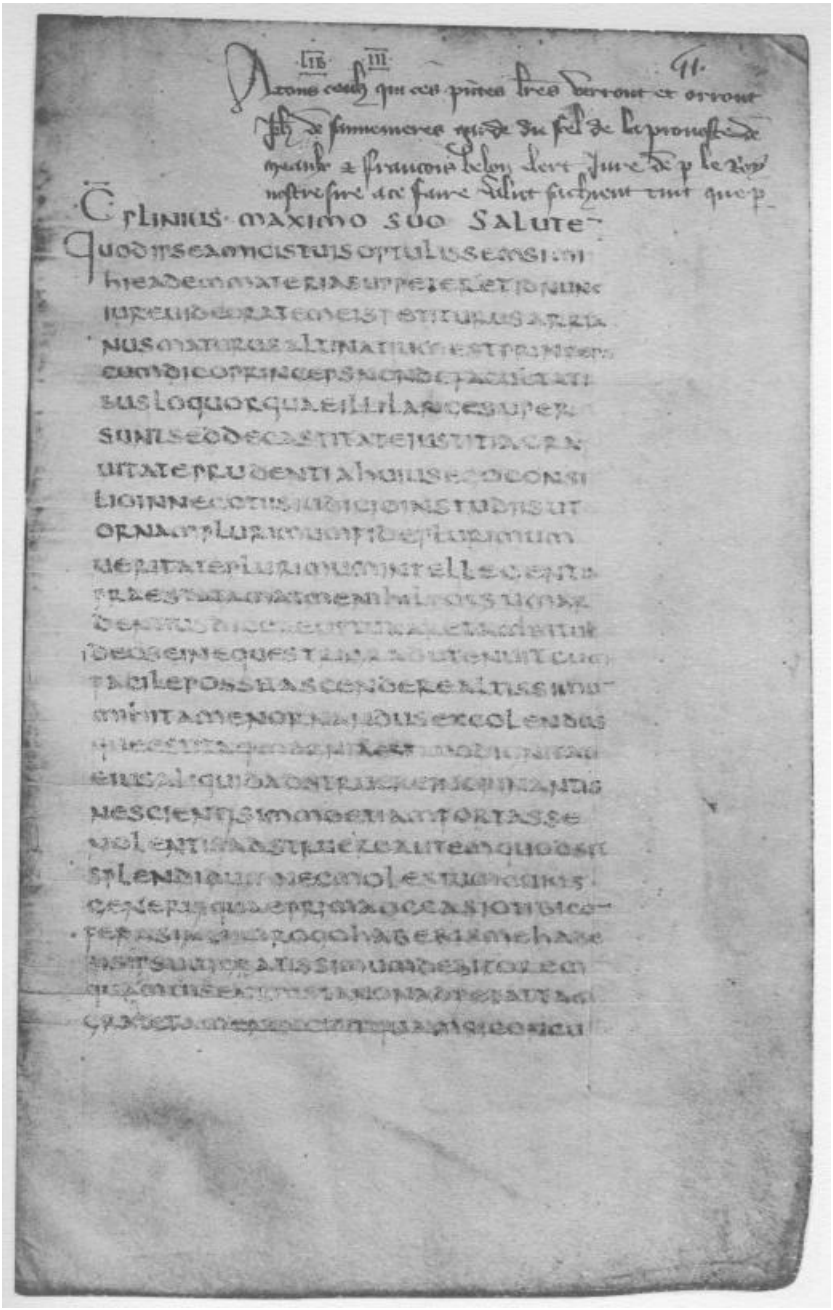


Fig. 5.2 Example of 6th Century *scriptura continua* (Lowe and Rand, 1922)

The introduction of spaces enabled efficient silent reading. Without spaces, there are no easily-perceptible units for the eyes to saccade to (see section 5.2.1). This is not even considering the many subsequent innovations in the development of line and page numbers, verse, paragraphs, chapters, volumes, indices, contents pages and guides for the religious and secular reader alike that enabled a silent reader to select any required section of text. Spaces that create the units of words allow a system of dual focus at the centre

and periphery which enables the development of units of reading. Identifying various units and joining them into predictable patterns produces easily-identified patterns of high and low 'information' (in Shannon's (1948; 1950/1993) mathematical sense of entropy). It is most likely that silent reading depends on these patterns to chunk minimal units into larger units of written information. It is also quite possible that these minimal and larger information units contribute to Sinclair and Mauranen's proposed Linear Unit Grammar (2006) which is based on the concept of chunks of various sizes, but are typically based on the unit of group as in the current study.

The introduction of spaces did more than make the concept of word appear corporeal. The apparently trivial introduction of spaces between words created the opportunity for greater participation in reading, as well as individual, even subversive, interpretation of texts because the practice of reading was no longer necessarily a group activity, entailing the social sanctioning of hegemonic and dogmatic interpretations. The change from collective to individual reading habits in 14th-17th Century European history produced the conditions for subversive interpretations of texts that contributed to major social change (Saenger, 1997). A further consequence of introducing spaces between words was a range of developments in the grammar of the written language. One, perhaps surprising, change was in the degree of inflection first in written and then in spoken vernaculars. The increasing importance of words and their order started to replace earlier grammatical patterns based on inflection:

> The various conventions of word order and the word separation that characterized the written vernacular after 1200 encouraged the dropping of inflection that had aided the ancient reader of *sciptura continua* to recognize properly and accent words and to identify their grammatical roles, a function that was no longer necessary in the separated vernacular texts (Saenger, 1997, p.266)

It seems quite incredible that such a small change in the written language could have such a major impact on the structure of spoken and written language, but it seems likely that spoken and written modes have influenced each other throughout history.

With the apparently trivial technological innovation of spaces between words, written language was no longer required to conform to the same units of information as spoken language because it was no longer limited by the physical system of articulation. Instead, information structure in written English became constrained by the visual system, responding to the physical pressures of saccades and other features of the reading process (see section 5.2.1). Spaces enabled an independent realisation for written information structure by making units of written language visible to the saccading eye, and freeing

written language from the physical constraints of an intonation contour. However, while the system of spaced, written English could produce the elements required for units of information, it did not yet allow points of prominence – New and Given information.

5.2.3 Punctuation

It would appear, then, that spaces between words enabled the establishment of silent reading as the norm. In conjunction with the development of spaced words was the development of punctuation marks. This section will investigate how spaces and punctuation may combine to realise information structure in written English.

Punctuation is a frequent topic for prescribers of language style and taste. From Lowth (1762) to Truss (2003), there has been no end of advice on the best way to punctuate. Typically, punctuation in English has come to be seen as a balance between the syntactic and prosodic features of language. Prescriptivists of the syntactic school insist that punctuation derives from "logical" rules of grammar, while those of the prosodic school encourage writers to read their sentences aloud, listening to intonation and pausing in order to identify the correct positions for punctuation marks. Baron (2001) charts the historical influences of syntax and rhetoric on punctuation, predicting that the current influence of spoken language on electronic text-types such as e-mail may produce ever-shorter sentences in all written genres. Nunberg (1990), however, argues against comparing punctuation to its intonational equivalent, particularly when most written English will never be spoken, concluding that the comparison provides "a theoretically uninteresting account of what is in any event a not very good correlation." (p.15)

As well as identifying three functions for punctuation marks (the marking of a boundary, of status, or of a relation), Halliday (1989) conforms to the view that there are two influences on punctuation choice: "punctuation according to grammar, and punctuation according to phonology" (p.37). I would argue that although the latter style of punctuation may result in a text that transposes to spoken English easily – a text that is written to be spoken – the other style of punctuation results in a text that is not easy to read aloud – it is written to be read (Gregory and Carroll, 1978). A written-to-be-spoken text reflects patterns of speech, and is generally divided into units determined by the constraints of the articulatory system. In contrast, a written-to-be-read text is constrained by the visual system. The two systems permit distinct units of information that are likely to vary with the Genre and Register of written English (see Fig. 5.1).

The point about spaces and punctuation is not that they are designed to help the reader to speak the sentences aloud – unless that is the function of the written text – but that they are designed to help the silent reader read more efficiently. The introduction of spaces to written text created units of words which allowed the eyes to track, or saccade, across stretches of written text, enabling silent reading (Saenger 1997). However, the units of words created by spaces remain undifferentiated in the stream of continuous text. The textual metafunction requires a system that allows these units to be differentiated by value. While spaces create the units of words, it is likely that punctuation provides points of prominence to create units of textual information. It is likely, too, that high frequency functional words can also be identified easily on a saccade, and are passed over in the same way as punctuation marks. Words like *so, and, or, but,* and other common conjunctions which divide clauses are typically shown in eye-movement studies to be words that do not receive attention in a fixation (Rayner, 1998).

Punctuation marks (and high frequency functional conjunctions) create visual points that the silent reader can easily saccade to. Therefore, this seems a fairly 'natural' position to place the prominent, or newsworthy, item in the information unit, in the same way that it is 'natural' to place the newsworthy item in spoken English at the most audibly prominent point. Returning to psychological studies we find evidence that the end-of-clause and end-of-clause-complex positions have a significant influence on eye movement when reading. Rayner *et al.* (2000; Hirotani *et al.*, 2006) record a significantly lengthened fixation in places immediately preceding commas and full-stops. In comparison to the same syntactic information presented without commas, they note that although the fixation time is longer in versions with punctuation marks, overall reading speed is faster, helped by the subsequent saccade going further into the next clause than in clauses without punctuation. Although Rayner *et al.* (2000) use these findings to support the view that there is a clause 'wrap-up' effect – that readers resolve all potential ambiguities before proceeding to the next clause – the data emphasises the importance of the final position in a clause. When reading, our eyes saccade to punctuation marks not in order to read them, but in order to see what is prior to the mark. The visible mark has little textual meaning on its own, but points the reader to the preceding units which carry the textual value of an informational culmination in prominence[7].

[7] These studies also suggest that punctuation marks allow the reader to saccade beyond more thematic material than in text without punctuation, suggesting that punctuation assists peripheral vision in identifying both Information and Theme.

Units of information in silent reading are based on the visual system. The visual system chunks information, whether the information is objects or words (Magnuson *et al.* 1998; Conklin and Schmitt, 2008). When reading in silence, we no longer need to draw breath to articulate (see section 5.2.1), enabling greater spans for each unit than the limits imposed by the physical system of articulation. (Intonation units in speech average approximately 6.5 orthographic words, while there is an average of about 9.6 words between punctuation marks for the same material (Chafe, 1988)). Thus, a major function of punctuation is to chunk written text into suitably-sized units for a silent reader to process more efficiently than they can speak. These units in written English need bear little relation to equivalent units in spoken English, unless they are intended to mirror spoken language. Written text that is intended to be read aloud will be shaped by the constraints of both the visual and articulatory systems, while text that will not be read aloud need only take the visual system into account. In this perspective, a simple sentence can be defined as a unit of Information for the written mode.

To summarise, punctuation marks probably perform the same informational function as the intonation contour in spoken English, but they are not constrained by the same physical systems that determine spoken language. Punctuation is able to exploit the potentialities of the human visual, rather than articulatory, system. This explains why there is no necessary direct correlation of the intonation pattern in the loud reading of a sentence and its punctuation. It also suggests that punctuating a sentence according to how it is said will not produce the most efficient writing as far as the silent reader is concerned. The systems of information structure in the two modes of speech and writing are functionally identical but respond to a different set of constraints (see section 5.1). Spoken language achieves prominence in information structure through the tonic foot, which is easily distinguished by the ear. Written language is hypothesised here to achieve prominence in information structure through a boundary mark that is easily recognisable to the saccading eye – typically a boundary-marking punctuation mark or a common function word to divide clauses. That is, punctuation marks are easily recognised by the visual system. They provide a means of dividing the stream of written language and selecting which items will be given prominence.

The system of written English which has enabled silent reading has implications for writers of English. Written language benefits from being produced in a time frame allowing for conscious choices and redrafting. Conscious attention and time allows

written language to be manipulated and changed more than spoken language. Thus, clausal elements can be manipulated through features such as clefts and passive voice, ensuring that the writer has control over what appears in final position as New information (Hannay and Martínez Caro, 2008). It is for this reason that I prefer the term 'special informational structures' over 'special thematic structures' as it is the system of Information more than the thematic system that motivates these choices. It is often noted that passive voice is more frequently associated with written than spoken language (Biber *et al.* 1999). It is likely that this is another consequence of demanding that sequence realises Information Structure in written language, rather than the more flexible assigning of the tonic foot to New information in spoken language (see section 4.7). The following section examines wider implications in the study of English.

5.3 Implications

This seems an appropriate point to summarise the model that is being proposed. Information structure functions neither to indicate what is referentially new in a text – that is the role of Presenting reference and other systems – nor to say something about what is being discussed in a Topic-Comment style sequence – that is one of the roles of Theme and Rheme. Information structure functions in text to divide experiential and ideational elements into units which contain New information. New information is that part of the message that the speaker or writer directs the listener's or reader's attention to; it is the *newsworthy* part of the message. In English, intonation divides spoken language into units of information, and New information is realised by the tonic foot. There is an unmarked correlation between a spoken clause and an information unit, but they are easily separated by intonation. In English, punctuation divides written language into units of information, and New information is realised by the constituent that is placed before a punctuation mark or other easily-visible sign to indicate prominence. There is an unmarked correlation between a written clause and an information unit, but they are easily separated by punctuation. The distinct features in the spoken and written modes derive from the constraints and affordances offered by the respective phonological and graphological systems of realisation in English.

The preceding discussion raises a number of questions that drive to the heart of core concepts in many linguistic models. These issues are discussed in the following sections

relating to, respectively, the status of the word as a unit and the status of the sentence as independent of the written mode.

A wide range of grammatical analyses have assumed *a priori* the linguistic units of the word and the sentence. It seems, however, that these units may be the consequence of a particular writing system, rather than a natural part of all languages. The gaps between sentences and the spaces between words are the consequence of literacy and of the view of language from a literate society. If a sentence, signalled by punctuation marks, functions to create gaps between ideas we must explain why there appear to be no clear relationships between punctuation marks and pauses (Davies, 1986). It seems that we need literacy to analyse language, but that literacy obscures the nature of language. This dilemma is most likely the source of the written bias in linguistics (Linell, 2005).

5.3.1 The Word as an Arbitrary Linguistic Unit

There is evidence to suggest that putting spaces between words is a comparatively recent innovation in the alphabetic systems of written Latin and modern European languages (Saenger, 1982; 1997). Dividing the stream of written language using spaces has, in effect, produced the category of word – without spaces there are no words since it is the space that defines a word. Without spaces, there are no words, and without written language a culture rarely develops a word for *word*: "Languages that exist in oral form only do not have a word for "word" because they do not have a conception of the linguistic unit that constitutes a word." (Saenger 1997 p.253). Phylogenetically, spaces and words are a late development in some languages. Although we may presume a principled methodology in the development of spacing words, how exactly words are separated varies between languages and is one area of research (*e.g.* Tuttle, 2008). Ontogenetically, too, literacy plays a central role in the developing concept of word. Bialystok (1986) and Roberts (1992) note how the implicit and explicit understanding of the word *word* is intimately tied to developing literacy. The centrality of the word to some contemporary linguistic models is undermined when its arbitrariness is considered; there is no unit of word identifiable in the stream of spoken language. It is quite possible that our view of language as written has affected the way we think of language to make us believe that the word is a natural, not arbitrary, unit. Evidence for this rather controversial perspective can be found in both a historical and a cross-linguistic comparison of speech and writing.

Historically, since the *scriptura continua* of the Latin scripts in medieval Europe, various systems of writing have evolved to divide words (Saenger 1997), including use of midline points between words and placing spaces at almost unpredictable places (see examples 5.2 and 5.3, respectively).

5.2 use♦of♦midline♦points♦between♦words

5.3 an dpl acing spacesatal most unpr edictab le pla ce s.

While a number of systems did not survive because they offered little improvement over uninterrupted text, one system that did endure featured the combination of single-syllable adpositions being attached to their nominal objects (Saenger 1997). While this practice may have stopped in English and French, it can be seen in the written systems of other languages, including Turkish. It is typical for Turkish to combine prepositions with their objects so that *at home* becomes *evde* (ev-de: home-at), *in the future* becomes *gelecekte* (gelecek-te: future-in) and *of the optician* becomes *gözlükcüsü* (göz-lük-cü-sü: eye-ness-er-of).

Although the manifestation of written words on a page give them a concrete appearance, their independence in speech is extremely difficult to verify: "The exact definition of units like words is not a given, self-evident fact of spoken language." (Linell, 2005, p.14) Certainly it seems that certain combinations of sounds recur throughout the language, but this fact alone would result in a syllabic-based writing system and linguistics. Halliday (1989) points out that pre-literate societies rarely develop a metalanguage. The concept of sentence, word, and grammar are subsequent to, and generally dependent on, the written form of a language. When language is written down it can be analysed far more easily, which often means that it is the written language that is theorised. Herein lies the dilemma for much of linguistics: it is far easier to study language when it is written down, but when language is written down it offers a distorted view of spoken language. The effect of the written language on much of linguistics has been to emphasise the units of sentence and word and consider them both immutable, without considering that these units, when defined by spaces and punctuation, may be historical language-specific phenomena developed for the convenience of the silent reader, not necessarily natural divisions of language.

5.3.2 The Sentence as a Convenience for the Reader

If we return to Halliday's (1967a; 1967b; 1976; 1994) characterisation of information structure in spoken language, it is pertinent that speech is always described as being analysable into units of information independently of clause grammar. While there is an unmarked correlation between Information Unit and clause, (and between New information and Presenting reference (Fries, 2000)), there is no defining relationship. The units of analysis and their realisations are independent. Spoken language, particularly unplanned exploratory or sharing dialogue (such as a discussion or conversation, see Fig. 5.1), contains almost as many non-clauses as clauses; interjections, false-starts, phrases, groups, and non-finite and incomplete clauses mix with full, grammatically-analysable clauses (Carter and McCarthy, 1997). However, while some spoken English cannot be analysed into clauses, all spoken English can be analysed into units of Information; information structure is a consistent unit of analysis in spoken language, regardless of 'grammaticality' according to a clausal or sententially-defined grammar. That is, intonation is a necessary part of the grammar of spoken language, realising information structure, while clause structure is not. While written language characteristically contains more consistent clause structure than spoken language, clause structure is probably not as central to silent reading as punctuation, which functions to realise information units that are distinct from those in the spoken mode. Clausal grammar produces typical written structures, but these can, and often are, superseded by choices in punctuation; punctuation appears to be more influential than clausal structure in dividing discourse into units of messages, just as intonation supersedes clausal structure in spoken language.

Section 5.2.3 discussed the development of punctuation. As spaces were introduced between words, and reading became a silent activity carried out by individuals, punctuation allowed the silent reader to interpret text by dividing spaced words into units, by showing logical relations, or by showing the speech function of a unit (Halliday, 1989). Punctuation, in all these cases, replaced the role of phonological or grammatical markers in texts that were previously read aloud in order to divide language into units of information based on spoken language. From this perspective one could suggest that information structure in writing is a product of the sentence. Alternatively it would seem more logical, taking a developmental perspective, to propose that the orthographic sentence is the product of information structure in writing. The sentence provides a visual method of dividing or chunking written discourse into distinct units and selecting elements for prominence. It may be the case that this is not the only factor involved in the

development of the unit of a sentence, but it is one that has been largely overlooked and, I am proposing, is a primary motivation. While the placing of divisions in written English has become largely determined by grammatical choices, many of these choices themselves developed as a result of the interaction of spaced writing, punctuation and speech (Saenger, 1997). The function of information structure is the same in spoken and written language, but is realised by intonation and punctuation, respectively. As written English has failed to develop a system for intonation we can conclude that the resources available are sufficient to realise the function of information structure; specifically, word spaces and punctuation suffice to divide written language into manageable chunks because they exploit the visual rather than the articulatory system.

Assuming the argument as presented so far, we must accept a very significant implication for the study of syntax, where syntax is limited to an orthographic sentence. If sentences do not exist except in the scripting of a language, then sentence-based syntax is a product of literate societies, and not necessarily that of the mind. That is, if information structure is the primary functional motivation for the unit of a sentence, and therefore constrains clausal grammar into units convenient to a silent reader, then the sentence has no definition independent of written information structure, and is a product of demands for an orthographic system that can respond to the demands of silent reading. Certainly, spoken language exhibits far too many cases of non-sentential utterances to be ignored as a major feature of language. The strict version of this hypothesis would dictate that the sentence exists only to realise information structure – it has no other definition or function. Thus, if the domain of syntax is the sentence, and the domain of information structure is the sentence then, as Fronek (1983) describes, information structure and syntax must indeed be intricately linked. This would also explain why theorists such as Steedman, Jackendoff, and Lambrecht can only identify information structure in terms of sentence structure; because they take the unit of the sentence as given and do not account for how information structure makes meaning in spontaneous spoken English without sentential grammar.

I am not arguing that clausal grammar does not exist – either in speech or in writing – as language evidently can be analysed as described in the numerous models that have been developed, although spoken language often expresses meanings that do not conform to clausal grammar as closely as written language. What I am arguing, however, is that the unit of the sentence is a graphological unit (Halliday and Matthiessen, 2004). Grammatical models that take the sentence as the basic unit of analysis can only describe the language

that we see on the page, because the sentence is not a grammatical unit of spoken language, but is a unit developed for the convenience of the silent reader. The sentence employs spaces and punctuation marks to divide the stream of language into analysable chunks, probably allowing for the eye to combine features of punctuation with group structure within a probabilistic learning model to allow peripheral vision to saccade to the next potential focus of information.

5.4 Further Research

Clearly, the conclusions reached here can only be presented provisionally. A great deal more research, some of which has been described at the end of each chapter, needs to be completed to verify the claims. The study here used only one register in its text analysis. A larger corpus with a wider range of texts, comparing across Registers, would help to identify the applicability of the theory to a range of written texts. A more consistent model that includes the effect of the logical metafunction would also improve the approach, and the analysis could also be made more detailed, or at a finer level of delicacy, by looking into the structure of the groups within the clause to see if similar effects operate at these other ranks.

Particularly needed at this point is empirical data that can reveal the nature of information structure in written English. One such experiment could manipulate the sequence of groups within a clause and both gauge readers' reactions to the readability of the text and time the readers in completing the task of reading. Eye-movement studies might also help in studies of manipulated texts. It is hypothesised that more frequent regression and shorter saccades would result from the disruption of the flow of text against the expected, or unmarked, pattern of conflating (Given) Theme with Presuming reference and (New) Final Rheme with Presenting reference, as identified in this study.

Further evidence may also be found in a historical approach. It is hypothesised that Special Informational structures (see section 4.7), such as cleft and pseudo-cleft structures, have increased in frequency and variety with the development of silent reading and increased literacy. Ball (1994) and Traugott (2007) provide preliminary support for this hypothesis.

Finally, the applications of the results of this study also need to be studied. It is hoped that teaching the model of Information Structure to second language readers – those that are expected to read the texts in this survey – will enable them to identify the main points of information and to recognise the flow of ideas in written discourse more readily. It is expected that tools such as the graphical reworking of a text, as illustrated in Fig. 4.7, would assist greatly in this application.

5.5 Conclusion

What we understand today as a sentence in the languages of western Europe is the result of a socio-historical process brought about by a particular set of historical moments, probably originating in the need by Irish monks to reproduce Latin texts (Saenger, 1982; 1997). It was only with the systems of spaces and punctuation marks that words and sentences came into existence, and it was only with the technology of words and sentences that we were able to invent the practice of silent reading. As silent reading developed, written language was required to realise many of the same linguistic functions as spoken language. A fundamental function performed by the textual metafunction in spoken language is to divide the stream of language into manageable chunks which are then differentiated between prominent and less prominent parts. This study has proposed, through discourse analysis and then through argumentation from a variety of sources, that this function is also carried out in written language, and that this is one of the main roles demanded of spaces and punctuation (Hannay and Kroon, 2005).

Since written language has not evolved a system of intonation to mirror the same system as speech for realising information structure, and information structure appears to operate in written language, written language needed a different system of realisation for information structure. It is this function that the sentence, defined as a graphological convention signified by punctuation marks, has evolved to fulfil. However, the written sentence is not constrained by the physical system of vocal production because the written mode allows the reader to 'voice' the words virtually, without exercising the articulatory organs. Consequently, the sentence becomes a different unit to the spoken information unit, and the only time that they coincide is when written language is produced to mirror the characteristics of speech. Both spoken and written English show an unmarked correlation between an information unit and a clause in the same mode, but

vary across modes; a spoken information unit rarely correlates with a written information unit. Silent reading is possible because spaces allow the writer to divide the stream of language into units which are then distinguished in prominence by punctuation and other visual marks. Information structure in written English is realised by spaces and punctuation.

References

Almor, A. 1999. Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review 106/4* p.748-765

Arbib, M. 2000. The mirror system, imitation and language. In C. Nehaniv and K. Dautenhahn (eds.), *Imitation in Animals and Artefacts*. Cambridge: MIT Press

Arnold, J. 2001. The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes 31/2* p.137-162

Arnold, J., Wasow, T., Losongco, A. and Ginstrom, R. 2000. Heaviness vs. Newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76/1 p.28-55

Asher, N. and Lascarides, A. 1998. Bridging. *Journal of Semantics 15* p.83-113

Banks, D. 2004. Degrees of newness. In D. Banks (ed.), *Text and Texture – Systemic Functional Viewpoints on the Nature and Structure of Text*. Paris: L'Harmattan

Baron, N. 2001. Commas and canaries: The role of punctuation in speech and writing. *Language Sciences 23* p.15-67

Barsalou, L.W. 2008. Grounded cognition. *Annual Review of Psychology 59* p.11.1-11.29

Behrens, H. 2009. Usage-based and emergentist approaches to language acquisition. *Linguistics 47/2* p.383-41

Benson, J.D and Greaves, W.S. (eds.) 2005. *Functional Dimensions of Ape-Human Discourse*. London: Continuum

Berry, M. 1995. Thematic options and success in writing. In M. Ghadessy (ed.), *Thematic Development in English Texts*. London: Pinter

Berry, M. 1996. What is Theme? A(nother) personal view. In M. Berry, C.S. Butler, R. Fawcett and G. Huang (eds.), *Meaning and Form*. New Jersey: Ablex

Bialystok, E. 1986. Children's concept of word. *Journal of Psycholinguistic Research 15/1* p.13-32

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

Birner, B.J. 1994. Information status and word order: An analysis of English inversion. *Language 70/2* p.233-259

Black, B. J. 1997. *Workshop Practices and Materials*. London: Butterworth-Heinemann

Bloor, M. 1983. The pragmatics of word omission in abbreviated texts: The case of encyclopedia entries. *MALS Journal New Series vol.8* p.105-123

Bloor, M. 1998. Lexical and grammatical choices in innovative language use in computer science. In A. Sanchez-Macarro and R. Carter (eds.), *Linguistic Choice across cultures: Variation in Spoken and Written English*. Amsterdam: John Benjamins

Bloor, M., and Bloor, T. 1992. Given and New information in the thematic organization of text: An application to the teaching of academic writing. *Occasional Papers in Systemic Linguistics no.6* p.33-43

Bolinger, D.L. 1958. Stress and information. *American Speech 33/1* p.5-20

Bolinger, D.L. 1965. Forms of English. Cambridge, MA: Harvard University Press

Bos, J. 2003. Implementing the binding and accommodation theory for anaphora resolution and presupposition projection. *Computational Linguistics 29/2* p.179-210

Bowcher, W. 2004. Theme and New in play-by-play radio sports commentating. In D. Banks (ed.), *Text and Texture*. Paris: L'Harmattan

Bransford, J. and Johnson, M. 1973. Considerations of some problems of comprehension. In W.G. Chase (ed.), *Visual Information Processing*. New York: Academic Press

Briggs, C.F. 2000. Literacy, reading and writing in the medieval West. *Journal of Medieval History 26/4* p.397-420

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

Brysbaert, M. and Praet, C. 1992. Reading isolated words: No evidence for automatic incorporation of the phonetic code. *Psychological Research 54* p.91-102

Buchla, D. and McLachlan, W. 1992. *Applied Electronic Instrumentation and Measurement*. Englewood Cliffs: Prentice Hall

Butler, C. 1985. *Statistics in Linguistics*. Oxford: Blackwell. Retrieved from http://www.uwe.ac.uk/hlss/llas/statistics-in-linguistics/bkindex.shtml on 6 Feb 2009

Butler, C.S. 2003. *Structure and Function Part 2 - From Clause to Discourse and Beyond*. Amsterdam: John Benjamins

Butler, C.S. 2005. Focusing on focus: A comparison of Functional Grammar, Role Reference Grammar and Systemic Functional Grammar. *Language Sciences 27* p.585-618

Butler, C.S. 2008. Cognitive adequacy in structural-functional theories of language. *Language Sciences 30/1* p.1-30

Bybee, J. 2006. From usage to grammar: The mind's response to repetition. *Language* 82/4 p.711-733

Caffarel, A. 2000. Interpreting French theme as a bi-layered structure: Discourse implications. In Ventola, E. (ed.), *Discourse and Community*. Tubingen: Gunter Narr Verlag

Caffarel, A., Martin, J.R. and Matthiessen, C.M.I.M. (eds.) 2004. *Language Typology. A Functional Perspective*. Amsterdam: John Benjamins

Caplan, D. 2004. Functional neuroimaging studies of written sentence comprehension. *Scientific Studies of Reading 8/3* p.225-240

Carter, R.A. and McCarthy, M.J. 1997. *Exploring Spoken English*. Cambridge: Cambridge University Press

Chafe, W. 1970. *Meaning and the Structure of Language*. Chicago: University of Chicago Press

Chafe, W. 1974. Language and consciousness. *Language 50/1* p.111-133

Chafe, W. 1976. Givenness, contrastiveness, definiteness, subjects and topics. In C. Li (ed.), *Subject and Topic*. New York: Academic Press

Chafe, W. 1988. Punctuation and the prosody of written language. *Written Communication 5/4* p.395-426

Chafe, W. 1991. Grammatical subjects in speaking and writing. *Text 11/1* p.45-72

Chafe, W. 1995. Accessing the mind through language. In S. Allén (ed.), *Of Thoughts and Words - Proceedings of Nobel Symposium '92*. London: Imperial College Press

Chater, N. and Manning, C.D. 2005. Probabilistic models of language processing and acquisition. *TRENDS in Cognitive Science 10/7* p.335-344

Clark, H.H. 1977. Bridging. In P.N. Johnson-Laird and P.C Wason (eds.), *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press

Clark, H.H. and Haviland, S. 1977. Comprehension and the Given-New contract. In R. Freedle (ed.), *Discourse Production and Comprehension*. New Jersey: Ablex

Clark, H.H. and Marshall, C. 1981. Definite reference and mutual knowledge. In A. Joshi, B. Webber and I. Sag (eds.), *Elements of Discourse Understanding*. Cambridge: Cambridge University Press

Coffin, C. 1997. Constructing and giving value to the past: An investigation into secondary school history. In F. Christie and J.R. Martin (eds.), *Genre and Institutions – Social Processes in the Workplace and School*. London: Cassell

Cognitive Science Laboratory, Princeton 2005. *WordNet*. Retrieved from http://wordnet.princeton.edu on 2 June 2005

Collins, P. 1991. *Cleft and Pseudo-Cleft Constructions in English*. London: Routledge

Coltheart, M. and Coltheart, V. 1997. Reading comprehension is not exclusively reliant upon phonological representation. *Cognitive Neuropsychology 14/1* p.167-175

Conklin, K. and Schmitt, N. 2008. Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics 29/1 p.72-89*

Coope, S., Cowley, J. and Willis, N. 2002. *Computer Systems: Architecture, Networks and Communications*. London: McGraw Hill

Coulouris, G., Dollimore, J. and Kindberg, T. 2001. *Distributed Systems: Concepts and Design. 3rd Edition*. Harlow: Pearson

Crompton, P. 2004. Theme in discourse: 'Thematic progression' and 'method of development' re-evaluated. *Functions of Language 11/2* p.213-49

Cummings, M. 2000. The inference of given information in written text. In E. Ventola (ed.), *Discourse and Community*. Tubingen: Gunter Narr Verlag

Cummings, M. 2005. The role of Theme and Rheme in contrasting methods of organization in texts. In C.S. Butler, M. Gómez-González and S.M. Doval-Suarez (eds.), *The Dynamics of Language Use*. Amsterdam: John Benjamins

Cutting, L.E. *et al*. 2006. Differential components of sentence comprehension: Beyond single word reading and memory. *NeuroImage* 29 p.429-438

Daneš, F. 1974. Functional sentence perspective and the organization of the text. In F. Daneš (ed.), *Papers on Functional Sentence Perspective*. Prague: Academia

Daneš, F. 1989. 'Functional sentence perspective' and discourse connectedness. In M. Conte, J.S. Petöfi and E. Sözer (eds.), *Text and Discourse Connectedness*. Amsterdam: John Benjamins

D'Ausillio, A. *et al*. 2009. The motor somatotopy of speech perception. *Current Biology 19/5* p.381-385

Davidse, K. 1987. M.A.K. Halliday's functional grammar and the Prague school. In R. Dirven and V. Fried (eds.), *Functionalism in Linguistics*. Amsterdam: John Benjamins

Davies, M. 1986. Literacy and intonation. In B. Couture (ed.), *Functional Approaches to Writing*. London: Frances Pinter

Davies, M. 1989. Prosodic and non-prosodic cohesion in speech and writing. *Word 40/1-2* p.255-261

Davies, M. 1994a. "I'm sorry, I'll read that again": Information structure in writing. In S. Čmejrkova and F. Štícha (eds.), *The Syntax of Sentence and Text: A Festschrift for František Daneš*. Amsterdam: John Benjamins

Davies, M. 1994b. Intonation IS visible in written English. In S. Čmejrkova, F. Daneš, and E. Havlova (eds.), *Writing vs. Speaking: Language, Text, Discourse Communication*. Tübingen: Gunter Narr Verlag

Dehaene, S., Cohen, L., Sigman, M. and Vinikier, F. 2005. The neural code for written words: A proposal. *TRENDS in Cognitive Science 9/7* p.335-341

Dehé, N. 2002. *Particle Verbs in English: Syntax, Information Structure and Intonation.* Amsterdam: John Benjamins

den Dikken, M. 2005. A comment on the topic of topic-comment. *Lingua 115* p.691-710

Devlin, J.T. and Aydelott, J. 2009. Speech perception: Motoric contributions versus the motor theory. *Current Biology 19/5* p.198-199

Downing, A. 1991. An alternative approach to theme: a systemic-functional perspective. *Word 42/2* p.119-143

Dubois, B.L. 1987. A reformulation of thematic progression typology. *Text 7/2* p.89-116

Edelman, G.M. 1999. Building a picture of the brain. *Annals of the New York Academy of Sciences 882 June 1999* p.68-89

Edelman, G.M. 2004. *Wider Than the Sky - The Phenomenal Gift of Consciousness*. New Haven: Yale University Press

Eiler, M. 1986. Thematic distribution as a heuristic for written discourse function. In B. Couture (ed.), *Functional Approaches to Writing*. London: Frances Pinter

Ellis, N.C. 2006. Language acquisition as rational contingency learning. *Applied Linguistics* 27/1 p.1-24

Emmott, C. 1994. Frames of reference: Contextual monitoring and the interpretation of narrative discourse. In M. Coulthard (ed.), *Advances in Written Text Analysis*. London: Routledge

Emmott, C. 1997. *Narrative Comprehension - A Discourse Perspective*. Oxford: Clarendon Press

Fang Y., McDonald, E. and Cheng M. 1995. On Theme in Chinese: From clause to discourse. In R. Hasan and P.H. Fries (eds.), *On Subject And Theme.* Amsterdam: John Benjamins

Fawcett, R. 1988. The English personal pronouns: an exercise in linguistic theory. In J.D. Benson, M. Cummings & W.S. Greaves (eds.), *Linguistics in a Systemic Perspective.* Amsterdam: John Benjamins

Fawcett, R. 2000. *A Theory of Syntax for Systemic Functional Linguistics*. Amsterdam: John Benjamins

Fawcett, R. 2003. The many types of 'Theme' in English: their semantic systems and functional syntax. Retrieved on November 7, 2003 from http://www.wagsoft .com/Systemics/Archive/Fawcett_ThemePaper.rtf

Fiddick, L., Cosmides, L., and Tooby, J. 2000. No interpretation without representation: the role of domain-specific representations in the Wason selection task. *Cognition 77* p.1–79

Firbas, J. 1968. On the prosodic features of the modern English finite verb as means of Functional Sentence Perspective. *Brno Studies in English 7* p.11-48

Firbas, J. 1972. On the interplay of prosodic and non-prosodic means of functional sentence perspective. In V. Fried (ed.), *The Prague School of Linguistics and Language Teaching*. London: Oxford University Press

Firbas, J. 1986. On the dynamics of written communication in the light of the theory of functional sentence perspective. In C.R. Cooper and S. Greenbaum (eds.), *Studying Writing: Linguistic Approaches*. Beverly Hills: Sage

Firbas, J. 1987. On the delimitation of the Theme in functional sentence perspective. In R. Dirven and V. Fried (eds.), *Functionalism in Linguistics*. Amsterdam: John Benjamins

Firbas, J. 1992. *Functional Sentence Perspective in Written and Spoken Communication*. Cambridge: Cambridge University Press

Flowerdew, J. 2003. Signalling nouns in discourse. *English for Specific Purposes 22* p.329-346

Francis, G. 1986. *Anaphoric Nouns*. Birmingham: ELR, University of Birmingham

Francis, G 1989. Thematic selection and distribution in written discourse. *Word 40/1-2* p.201-221

Fries, P.H. 1981. On the status of Theme in English: Arguments from Discourse. *Forum Linguisticum 6/1* p.1-38

Fries, P.H. 1992. The structuring of information in written text. *Language Sciences 14/4* p.461-488

Fries, P.H. 1994. On theme, rheme and discourse goals. In M. Coulthard (ed.), *Advances in Written Text Analysis*. London: Routledge

Fries, P.H. 1995. Themes, Methods of development and texts. In R. Hasan and P.H. Fries (eds.), *On Subject And Theme*. Amsterdam: John Benjamins

Fries, P.H. 1995a. A personal view of theme. In M. Ghadessy (ed.), *Thematic Development in English Texts*. London: Pinter

Fries, P.H. 1995b. Patterns of information in initial position in English. In P.H. Fries and M. Gregory (eds.), *Discourse in Society: Systemic Functional Perspectives*. Norwood: Ablex

Fries, P.H. 1997. Theme and New in Written English. In T. Miller (ed.), *Functional Approaches to Written Text*. Washington: USIS

Fries, P.H. 2000. Issues in modelling the textual metafunction. In M. Scott and G. Thompson (eds.), *Patterns of Text: In honour of Michael Hoey*. Amsterdam: John Benjamins

Fries, P.H. 2002. The flow of information in a written text. In Fries, P., Cummings, M., Lockwood, D. and Spruiell, W. (eds.), *Relations and Functions Within and Around Language*. London: Continuum

Fries, P.H. 2009. The textual metafunction as a site for a discussion of the goals of linguistics and techniques of linguistic analysis. In G. Forey and G. Thompson (eds.) *Text Type and Texture*. London: Equinox

Fries, P.H. and Francis, G. 1992. Exploring Theme: Problems for research. *Occasional Papers in Systemic Linguistics 6* p.45-60

Froese, T. and Ziemke, T. 2009. Enactive artificial intelligence: Investigating the systemic organisation of life and mind. *Artificial Intelligence 173/3-4* p.466-500

Fronek, J. 1983. Some criticisms of Halliday's 'Information Systems'. *Lingua 60* p.311-329

Giora, R. 2002. Literal vs. figurative language: Different or equal? *Journal of Pragmatics 34/4* p.487-506

Girotto *et al.* 2001. Inept reasoners or pragmatic virtuosos? Relevance and the deontic selection task. *Cognition 81* p.B69-B76

Godelier, M. 1984. *The Mental and the Material*. London: Verso

Göksel, A. and Özsoy, A. 2003. dA: a focus/topic associated clitic in Turkish. *Lingua 113* p.1055-1088

Gómez, M. 1994. The relevance of Theme in the textual organization of BBC news reports. *Word 45/3* p.293-305

Gómez-González, M. 2001a. Some reflections on systemic functional grammar: with a focus on Theme. *Word 52/1* p.1-29

Gómez-González, M. 2001b. *The Theme-Topic Interface - Evidence from English*. Amsterdam: John Benjamins

Goody, J. 1986. *The Logic of Writing and the Organization of Society*. Cambridge: Cambridge University Press

Gosden, H. 1993. Discourse functions of subject in scientific research articles. *Applied Linguistics 14/1* p.56-75

Green, C.F., Christopher, E.R. and Mei, J.L.K. 2000. The incidence and effects of coherence of marked themes in interlanguage texts: A corpus-based inquiry. *English for Specific Purposes vol.19* p.99-113

Gregory, M. and Carroll, S. 1978. *Language and Situation: Language Varieties and their Social Contexts*. London: Routledge.

Grice, H.P. 1975. Logic and conversation. In P. Cole and J. Morgan (eds.), *Speech Acts*. New York: Academic Press

Grosz, B. 1977. The Representation and Use of Focus in Dialogue Understanding. *SRI Technical Note 151*. Menlo Park, California: SRI International

Grosz, B. 1978. Focus spaces: A representation of the focus of attention of a dialog. In D. Walker (ed.), *Understanding Spoken Language*. New York: North-Holland

Grosz, B., Joshi, A. and Weinstein, S. 1995. Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics 21/2* p.203-225

Grosz, B. and Sidner, C. 1998. Lost intuitions and forgotten intentions. In M. Walker, A. Joshi and E. Prince (eds.), *Centering in Discourse*. Oxford: Oxford University Press

Gundel, J. 1985. Shared knowledge and topicality. *Journal of Pragmatics 9* p.83-107

Gundel, J., Hedberg, N. and Zacharski, R. 1993. Cognitive status and the form of referring expressions in discourse. *Language 69/2* p.274-307

Gundel, J, Hegarty, M. and Borthen, K. 2003. Cognitive status, information structure, and pronominal reference to clausally introduced entities. *Journal of Logic, Language and Computers 12* p.281-299

Hahn, U. and Strube, M. 1997. Centering in-the-large: Computing referential discourse segments, *Proceedings of the Eighth Conference on European Chapter of the Association for Computational Linguistics, Madrid, Spain* p.104-111

Hajicova, E. and Sgall, P. 1988. Topic and focus of a sentence and the patterning of a text. In J. Petöfi (ed.), *Text and Discourse Constitution*. Berlin: Walter de Gruyter

Halliday, M.A.K. 1964/1981. Syntax and the Consumer. Originally published in *Monograph Series in Languages and Linguistics 17*. Washington: Georgetown University Press. Reprinted in M.A.K. Halliday and J.R. Martin (eds.), *Readings in Systemic Linguistics*. London: Batsford

Halliday, M.A.K. 1967a. Notes on transitivity and theme part 2. *Journal of Linguistics 3/2* p.199-244

Halliday, M.A.K. 1967b. *Intonation and Grammar in British English*. The Hague: Mouton

Halliday, M.A.K. 1974. The place of "functional sentence perspective" in the system of linguistic description. In F. Daneš (ed.), *Papers on Functional Sentence Perspective*. Prague: Academia

Halliday, M.A.K. 1975. *Learning How to Mean: Explorations in the Development of Language*. London: Edward Arnold

Halliday, M.A.K. 1976. Theme and information in the English clause. In G. Kress (ed.), *Halliday: System and Function in Language*. London: Oxford University Press

Halliday, M.A.K. 1979. Modes of meaning and modes of expression: Types of grammatical structure, and their determination by different semantic functions. In D.J. Allerton, E. Carney and W. Holcroft (eds.), *Function and Context in Linguistic Analysis*. Cambridge: Cambridge University Press

Halliday, M.A.K. 1981/2002. Text semantics and clause grammar: How is a text like a clause? In J. Webster (ed.), *On Grammar*. London: Continuum

Halliday, M.A.K. 1985. *An Introduction to Functional Grammar*. London: Edward Arnold

Halliday, M.A.K. 1985/2004. It's a fixed word order language is English. In J. Webster (ed.), *Studies in English Language*. London: Continuum

Halliday, M.A.K. 1989. *Spoken and Written Language*. Oxford: Oxford University Press

Halliday, M.A.K. 1991. Corpus studies and probabilistic grammar. In K. Aijmer and B. Altenberg (eds.), *English Corpus Linguistics*. Harlow: Longman

Halliday, M.A.K. 1993. Towards a language-based theory of learning. *Linguistics and Education 5* p.93-116

Halliday, M.A.K. 1994. *An Introduction to Functional Grammar - Second Edition*. London: Edward Arnold

Halliday, M.A.K. 1995. Fuzzy grammatics: A systemic functional approach to fuzziness in natural language. *Proceedings of 1995 IEEE International Conference on Fuzzy Systems*, Yokohama, Japan, March 20-24, 1995. Vol.1 p.9-26

Halliday, M.A.K. 2002. Text semantics and clause grammar: How is a text like a clause? In J. Webster (ed.), *On Grammar*. London: Continuum

Halliday, M. and Greaves, W.S. 2008. *Intonation in the Grammar of English*. London: Equinox

Halliday, M. and Hasan, R. 1976. *Cohesion in English*. London: Longman

Halliday, M. and Hasan, R. 1985. *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective*. Victoria: Deakin University Press

Halliday, M. and James, Z. 1993. A quantitative study of polarity and primary tense in the English finite clause. In J. Sinclair, M. Hoey, and G. Fox (eds.), *Techniques of Description: Spoken and Written Discourse*. London: Routledge

Halliday, M. and Martin, J. 1993. *Writing Science - Literacy and Discursive Power*. London: The Falmer Press

Halliday, M.A.K. and Matthiessen, C.M.I.M. 1999. *Construing Experience through Meaning: A Language-based Approach to Cognition*. London: Continuum

Halliday, M.A.K. and Matthiessen, C.M.I.M. 2004. *An Introduction to Functional Grammar – Third Edition*. London: Arnold

Hannay, M. and Kroon, C. 2005. Acts and the relationship between discourse and grammar. *Functions of Language 12/1* p.87-124

Hannay, M. and Martínez Caro, E.M. 2008. Last things first: A FDG approach to clause-final focus constituents in Spanish and English. In M.A. Gómez-González, J.L. MacKenzie and E. Gonzalez-Alvarez (eds.), *Languages and Cultures in Contrast: New Directions in Contrastive Linguistics*. Amsterdam: John Benjamins

Hartnett, C. 1995. The pit after the theme. In M. Ghadessy (ed.), *Thematic Development in English Texts*. London: Pinter

Hasan, R. 1984. Coherence and cohesive harmony. In J. Flood (ed.), *Understanding Reading Comprehension*. Newark: International Reading Association

Hasan, R. 1984/1996. Ways of saying: ways of meaning. In R. Fawcett et al. (eds.), *The Semiotics of Culture and Language*. London: Frances Pinter. Reprinted in C. Cloran, D. Butt and G. Williams (eds.), *Ways of Saying: Ways of Meaning - Selected Papers of Ruqaiya Hasan*. London: Cassell

Hasan, R. and Fries, P. (eds.) 1995. *On Subject And Theme*. Amsterdam: John Benjamins

Hasselgard, H. 2004. The role of multiple themes in cohesion. In K. Aijmer and A. Stenstrom (eds.), *Discourse Patterns in Spoken and Written Corpora*. Amsterdam: John Benjamins

Haviland, S. and Clark, H. 1974. What's new? Acquiring New information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior 13* p.512-21

Hawking, S. 1988. *A Brief History of Time*. London: Bantam Books.

Hirotani, M., Frazier, L. and Rayner, K. 2006. Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language 54* p.425-443

Hoey, M. 1986. The discourse colony: a preliminary study of a neglected discourse type. In M. Coulthard (ed.), *Talking about Text*. Birmingham: ELR

Hoey, M. 1991a. Another perspective on coherence and cohesive harmony. In E. Ventola (ed.), *Functional and Systemic Linguistics - Approaches and Uses*. Berlin: Mouton de Gruyter

Hoey, M. 1991b. *Patterns of Lexis in Text*. Oxford: Oxford University Press

Hoffman, B. 1994. Generating context-appropriate word orders in Turkish. *Proceedings of 7th International Generation Workshop*, Kennebunkport, Maine p.117-126. Retrieved 5 October, 2007 from http://acl.ldc.upenn.edu/W/W94/W94-0314.pdf

Hoffman, B. 1995. Integrating "Free" word order syntax and information structure. *Proceedings of 1995 Conference of the European Chapter of Association for Computational Linguistics*, Dublin, Ireland, p.245-252

Horowitz, P and Hill, W. 1989. *The Art of Electronics*. Cambridge: Cambridge University Press

Huang, G. 1996. Experiential enhanced theme in English. In M. Berry, C.S. Butler, R. Fawcett and G. Huang (eds.), *Meaning and Form*. New Jersey: Ablex

Huang, G. and Fawcett, R. 1996. A functional approach to two 'focussing' constructions in English and Chinese. *Language Sciences 18/1-2* p.179-194

Huddleston, R. 1988. Review: Constituency, Multi-Functionality and Grammaticalization in Halliday's Functional Grammar. *Journal of Linguistics 24/1* p.137-174

Hunston, S. and Thompson, G. (eds.) 2000. *Evaluation in Text*. Oxford: Oxford University Press

Hunt, K.W. 1965. Grammatical Structures Written at Three Grade Levels. NCTE Research Report No. 3. Champaign, IL: NCTE

Jackendoff, R. 2002. *Foundations of Language*. Oxford: Oxford University Press

Jacobs, J. 2001. The dimensions of topic-comment. *Linguistics 39/4* p.641-681

Jared, D., Levey, B.A and Rayner, K. 1999. The role of phonology in the activation of word meanings during reading: Evidence from proofreading and eye movements. *Journal of Experimental Psychology: General 128/3* p.219-264

Jones, D. and Elcock, J. 2001. *History and Theories of Psychology*. London: Arnold

Jones, S. 2002. *Antonymy – A Corpus-based Perspective*. London: Routledge.

Juhasz, B.J. and Rayner, K. 2003. Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory and Cognition 29/6* p.1312-1318

Just, M.A., Carpenter, P.A. and Woolley, J.D. 1980. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111/2* p.228-238

Kaltenbock, G. 2005. It-extraposition in English. *International Journal of Corpus Linguistics 10/2 p.119-159*

Kong, K. 2006. A taxonomy of the discourse relations between words and visuals. *Information Design Journal and Document Design 14/3* p.207-230

Kravchenko, A. 2002. Towards a bio-cognitive philosophy of language. *Perspectives - Journal for Interdisciplinary Work in the Humanities 1/4*. Retrieved from http://cogprints.org/4002/01/Cogphilosophy _Language.html on 12 March 2006

Kravchenko, A. 2006. Cognitive linguistics, biology of cognition and biosemiotics: Bridging the gaps. *Language Sciences 28* p.51-75

Kress, G. and van Leeuwen, T. 1996. *Reading Images - The Grammar of Visual Design*. London: Routledge

Lambrecht, K. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press

Lappin, S. and Leass, H. 1994. An algorithm for pronominal anaphora resolution. *Computational Linguistics 20/4* p.535-61

Lassen, I. 2003. *Accessibility and Acceptability in Technical Manuals*. Amsterdam: John Benjamins

Lassen, I. 2004. Ideological resources in biotechnology press releases: Patterns of Theme/Rheme and Given/New. In L. Young and C. Harrison (eds.), *Systemic Functional Linguistics and Critical Discourse Analysis*. London: Continuum

Lawrence, D.H. (1913/1999) *Sons and Lovers*. New York: Modern Library.

Leckie-Tarry, H. 1995. *Language and Context - A Functional Linguistic Theory of Register*. London: Pinter

Lee, Y-A, Binder, K.S., Kim, J-O, Pollatsek, A. and Rayner, K. 1999. Activation of phonological codes during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance 25/4* p.948-964

Lemke, J.L. 1998. Multiplying meaning: Visual and verbal semiotics in scientific text. In J.R. Martin and R. Veel (eds.), *Reading Science - Critical and Functional Perspectives on Discourses of Science*. London: Routledge

Lemke, J.L. 2000. Material sign processes and emerging ecosocial organization. In Andersen *et al.* (eds.), *Downward Causation*. Aarhus: Aarhus University Press

Levinson, S. 1996. Language and Space. *Annual Review of Anthropology 25* p.353-82

Linell, P. 2005. *The Written Language Bias in Linguistics*. London: Routledge

Lowe, E.A. and Rand, E.K. 1922. *A Sixth-Century Fragment of the Letters of Pliny the Younger*. Cambridge, Mass.: Cambridge University Press. Retrieved 6 August, 2008 from http://www.gutenberg.org/files/16706/16706-h/16706-h.htm

Lowth, R. 1762/1967 A Short Introduction to English Grammar. A. Miller and J. Dodsby: London. Reprinted by Scholar Press: Menston.

Lukatela, G., Eaton, T., Lee, C. and Turvey, M.T. 2001. Does visual word identification involve a sub-phonemic level? *Cognition 78* p.B41-B52

Lyons, J. 1977. *Semantics* (Vols. 1 and 2). Cambridge: Cambridge University Press

MacWhinney, B. 1998. Models of the emergence of language. *Annual Review of Psychology* 49 p.199-227

Magnuson, J.S., Bensinger, D.G., Hayhoe, M. and Ballard, D. 1998. Learning to form visual chunks: On the structure of visuo-spatial working memory. In M.A. Gernsbacher and S.J. Derry (eds.), *Proceedings of 20th Annual Conference of the Cognitive Science Society*. New Jersey: Lawrence Erlbaum.

Martin, J.R. 1989. *Factual Writing: Exploring and Challenging Social Reality*. Oxford: Oxford University Press

Martin, J.R. 1991. Nominalization in science and humanities: distilling knowledge and scaffolding text. In E. Ventola (ed.), *Functional and Systemic Linguistics - Approaches and Uses*. Berlin: Mouton de Gruyter

Martin, J.R. 1992. *English Text: System and Structure*. Amsterdam: John Benjamins

Martin, J.R. 1992b. Theme, method of development and existentiality: The price of reply. *Occasional Papers in Systemic Linguistics 6* p.147-184

Martin, J.R. 1995a. More than what the message is about: English Theme. In M. Ghadessy (ed.), *Thematic Development in English Texts*. London: Pinter

Martin, J.R. 1995b. Text and clause: fractal resonance. *Text 15/1* p.5-42

Martin, J.R. 2000. Close reading: Functional linguistics as a tool for critical discourse analysis. In L. Unsworth (ed.), *Researching Language in Schools and Communities*. London: Cassell

Martin, J.R. and Rose, D. 2002. *Working with Discourse*. London: Continuum

Martin, J.R. and White, P.R.R. 2005. *Language of Evaluation - Appraisal in English*. London: Palgrave Macmillan

Matsui, T. 1999. On the role of context in relevance-based accessibility ranking of candidate referents. In P. Bouquet et al (eds.) *Proceedings of Context'99*. Berlin: Springer-Verlag

Matthiessen, C.M.I.M. 1988. Representational issues in systemic functional grammar. In J.D. Benson and W.S. Greaves (eds.), *Systemic Functional Approaches to Discourse*. New Jersey: Ablex

Matthiessen, C.M.I.M. 1992. Interpreting the textual metafunction. In M. Davies and L. Ravelli (eds.), *Advances in Systemic Linguistics: Recent Theory and Practice*. London: Pinter

Matthiessen, C.M.I.M. 1995. THEME as an enabling resource in ideational 'knowledge' construction. In M. Ghadessy (ed.), *Thematic Development in English Texts*. London: Pinter

Matthiessen, C.M.I.M. 1995a. *Lexicogrammatical Cartography: English Systems*. Tokyo: International Language Science Publishers

Matthiessen, C.M.I.M. 1995b. Fuzziness construed in language: A linguistic perspective. *Proceedings of 1995 IEEE International Conference on Fuzzy Systems*, Yokohama, Japan, March 20-24, 1995. Vol.4 p.1871-1878

Matthiessen, C.M.I.M. 2006. Educating for advanced foreign language capacities: Exploring the meaning-making resources of languages systemic-functionally. In H. Byrnes (ed.), *Advanced Language Learning: The Contribution of Halliday and Vygotsky*. London: Continuum

Maturana, H.R. 1978. Biology of language: The epistemology of reality. In G.A. Miller & E. Lenneberg (eds.) *Psychology and Biology of Language and Thought*. New York: Academic Press

Maturana, H.R. and Varela, F.J. 1987. *The Tree of Knowledge – The Biological Roots of Human Understanding*. Boston: Shambhala

Mauranen, A. 1996. Discourse Competence - Evidence from thematic development in native and non-native texts. In E. Ventola and A. Mauranen (eds.), *Academic Writing - Intercultural and Textual Issues*. Amsterdam: John Benjamins

Mauranen, A. 1999. What sort of Theme is *there*? *Languages in Contrast 2/1* p.57-87

McCabe, A. 2004. Thematic progression patterns and text types in history textbooks. In D. Banks (ed.), *Text and Texture – Systemic Functional Viewpoints on the Nature and Structure of Text*. Paris: L'Harmattan

McCord, M. 1989a. Design of LMT: A prolog-based machine translation system. *Computational Linguistics 15* p.33-52

McCord, M. 1989b. A new version of the machine translation system LMT. *Literary and Linguistic Computing 4* p.218-229

McLuhan, M. 1970. The effect of the printed book on language. In E. Carpenter and M. McLuhan (eds.), *Explorations in Communication*. London: Jonathan Cape Ltd.

Mitkov, R. 2000. Pronoun resolution: The practical alternative. In S. Botley and T. McEnery (eds.), *Corpus-based and Computational Approaches to Discourse Anaphora*. Amsterdam: John Benjamins

Mitkov, R., Lappin, S. and Boguraev, B. 2001. Introduction to the special issue on computational anaphora resolution. *Computational Linguistics 27/4* p.473-477

Moore, N. 2006. Advanced language for intermediate learners: Language needs analysis and curriculum specification in English for academic purposes. In H. Byrnes (ed.), *Advanced Language Learning: The Contribution of Halliday and Vygotsky*. London: Continuum

Moore, N. 2008a. Bridging the metafunctions: Tracking participants through taxonomies. In C. Jones & E. Ventola (eds.) *From Language to Multimodality: New Developments in the Study of Ideational Meaning*. London: Equinox

Moore, N. 2008b. Validating a model of information structure in written English through a reading protocol. In E. Steiner and S. Neumann (eds.) *Data and Interpretation in Linguistic Analysis. Proceedings of The 19th European Systemic Functional Linguistics Conference and Workshop, Saarbrücken, Germany, 23-25 July, 2007*. Retrieved 12 January, 2009 from http://scidok.sulb.uni-saarland.de/volltexte/2008/1697/pdf/Moore_form.pdf

Navarretta, C. 2002. Combining Information Structure and Centering-based Models of Salience for Resolving Intersentential Pronominal Anaphora. Retrieved 20 November, 2003 from http://cst.dk/costanza/papers/daarc.pdf

Nesbitt, C. and Plum, G. 1988. Probabilities in a systemic-functional grammar: The clause complex in English. In R. Fawcett and D. Young (eds.), *New Developments in Systemic Linguistics Vol. 2: Theory and Application*. London: Frances Pinter

North, S. 2005. Disciplinary variation in the use of Theme in undergraduate essays. *Applied Linguistics 26/3* p.431-452

Nunberg, G. 1990. *The Linguistics of Punctuation*. Stanford, CA: CSLI.

O'Donnell, M. 1990. A dynamic model of exchange. *Word 41/3* p.293-328

O'Donnell, M. 1999. Context in dynamic modelling. In M. Ghadessy (ed.), *Text and Context in Functional Linguistics*. Amsterdam: John Benjamins

O'Donnell, M. 2003. *Systemic Coder v4.63* Software. Retrieved 20 November, 2003 from http://www.wagsoft.com/

O'Donnell, M. 2007-2009. *UAM Corpus Tool v1.29-v2.0*. Software. Retrieved 14 December, 2007 & 15 January 2010 from http://www.wagsoft.com/CorpusTool/download .html

O'Halloran, K.L. (ed.), 2004. *Multimodal Discourse Analysis*. London: Continuum

Ong, W.J. 1982. *Orality and Literacy: The Technologizing of the Word*. Routledge: London

O'Toole, M. 1994. *The Language of Displayed Art*. London: Leicester University Press.

Peereman, R., Content, A. and Bonin, P. 1998. Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory and Language 39* p.151-174

Perfetti, C.A. and Bolger, D.J. 2004. The brain might read that way. *Scientific Studies of Reading 8/3* p.293-304

Pickering, L. 2004. The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse. *English for Specific Purposes 23* p.19-43

Pike, K.L. 1959. Language as particle, wave, and field. *The Texas Quarterly 2/2* p.37-54

Pollatsek, A. and Rayner, K. 1982. Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance 8* p.817-833

Prince, E. 1978. A comparison of WH-clefts and *it*-clefts in discourse. *Language 54* p.883-906

Prince, E. 1981. Toward a taxonomy of given-new information. In P. Cole (ed.), *Radical Pragmatics*. New York: Academic Press

Prince, E. 1985. Fancy syntax and 'shared knowledge'. *Journal of Pragmatics 9* p.65-81

Prince, E. 1992. The ZPG letter: Subjects, definiteness, and information-status. In W. Mann and S.A. Thompson (eds.), *Discourse Description*. Amsterdam: John Benjamins

Pullum, G.K. and Scholz, B.C. 2002. Empirical assessment of poverty of stimulus arguments. *The Linguistic Review 19* p.9-50

Pulvermüller, F., Hauk, O., Nikulin, V.V., and Ilmoniemi, R.J. 2005. Functional links between motor and language systems. *European Journal of Neuroscience 21* p.793–797

Quirk, R. and Greenbaum, S. 1973. *A University Grammar of English* Harlow: Longman

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. 1972. *A Grammar of Contemporary English*. Harlow: Longman

RAF Training Manual. n.d. *Retardation Methods*. ref: RAF PTC CN 3787 1-1-6 06-528a/01/B50 1-1-7

Rappaport, T.S. 2002. *Wireless Communications – Principles and Practice*. Upper Saddle River: Prentice Hall

Ravelli, L. 1995. A dynamic perspective: Implications for metafunctional interaction and an understanding of theme. In R. Hasan and P.H. Fries (eds.), *On Subject and Theme: A Discourse Functional Perspective*. Amsterdam: John Benjamins

Ravelli, L. 2004. Signalling the organization of written texts: Hyper-Themes in management and history essays. In L.J. Ravelli and R.A. Ellis (eds.), *Analysing Academic Writing – Contextualized Frameworks*. London: Continuum

Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin 124/3* p.372-422

Rayner, K., Kambe, G. and Duffy, S.A. 2000. The effects of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology 53A/4* p.1061-1080

Rayner, K., Pollatsek, A. and Binder, K.S. 1998. Phonological codes and eye movements in reading. *Journal of Experimental Psychology: Learning, Memory and Cognition 24/2* p.476-497

Richards, G. 1996. *Putting Psychology in its Place*. London: Routledge

Riesco-Bernier, S. 2002. Tonicity and the status of information: A functional analysis of New information in pre-school teacher talk. Conference Paper presented at 29th International Systemic Functional Congress, Liverpool, July 2002

Riesco-Bernier, S. and Romero-Trillo, J. 2008. The acoustics of 'newness' and its pragmatic implications in classroom discourse. *Journal of Pragmatics 40*, p.1103-1116

Rizzolatti, G. and Craighero, L. 2004. The Mirror-Neuron System. *Annual Review of Neuroscience 27* p.169-192

Roberts, B. 1992. The Evolution of the Young Child's Concept of "Word" as a Unit of Spoken and Written Language. *Reading Research Quarterly 27/2*, pp. 124-138

Roe, P. 1998. WinATA - Aston Text Analyser Mark 2.0. Software licensed by Forwords Ltd.

Romero-Trillo, J. 1994. Ahm, ehm, you call it theme? … A thematic approach to spoken English. *Journal of Pragmatics 22* p.495-509

Rose, D. 2001. Some variations in Theme across languages. *Functions of Language 8/1* p.109-145

Rumelhart, D. 1980. Schemata: the building blocks of cognition. In R. Spiro, B. Bruse and W. Brewer (eds.), *Theoretical Issues in Reading Comprehension*. New Jersey: Lawrence Erlbaum Assocs.

Saenger, P. 1982. Silent reading: Its impact on late medieval script and society. *Viator 13* p.367-414

Saenger, P. 1997. *Space Between Words: The Origins of Silent Reading*. Stanford, CA: Stanford University Press

Salmelin, R. and Kujala, J. 2006. Neural representation of language: Activation versus long-range connectivity. *TRENDS in Cognitive Science 10/11* p.519-525

Sampson, G. 1997. *Educating Eve - The 'Language Instinct' Debate*. London: Continuum

Sgall, P. 1975. On the nature of topic and focus. In H. Ringbom (ed.), *Style and Text*. Stockholm: Sprakforlaget Skriptor

Sgall, P. 1987. Prague functionalism and topic vs. focus. In R. Dirven and V. Fried (eds.), *Functionalism in Linguistics*. Amsterdam: John Benjamins

Shannon, C. 1948. A mathematical theory of communication. *The Bell System Technical Journal 27* p.379-423 *and* p.623-656. Retrieved 21 May 2002 from http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf

Shannon, C. 1950/1993. Prediction and entropy of printed English. Reprinted in N.J.A. Sloane and A.D. Wyner (eds.), *Claude Elwood Shannon: Collected Papers*. New York: IEEE Press

Shebilske, W. and Reid, L. 1979. Reading eye movements, macro-structure and comprehension. In P.A. Kolers, M.E. Wrolstad and H. Bouma (eds.), *Processing of Visible Language*. New York: Plenum

Sidner, C. 1978. The use of focus as a tool for disambiguation of definite noun phrases. Retrieved 5 August, 2004 from http://delivery.acm.org/10.1145/810000/803158/p86sidner.pdf?key1=803158andkey2=1509271901andcoll=GUIDEanddl=ACMandCFID=25222196andCFTOKEN=53831658

Sidner, C. 1979. The role of focussing in interpretation of pronouns, *Proceedings of the 17th Conference on Association for Computational Linguistics*, La Jolla, California, p.77 - 78

Sidner, C. 1981. Focusing for interpretation of pronouns. *American Journal of Computational Linguistics 7/4* p.217-231

Sinclair, J. 1988. Compressed English. In M. Ghadessy (ed.), *Registers of Written English*. London: Pinter

Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press

Sinclair, J. 1993. Written discourse structure. In J. Sinclair, M. Hoey, and B. Fox (eds.), *Techniques of Description: Spoken and Written Discourse*. London: Routledge

Sinclair, J. 1994. Trust the text. In M. Coulthard (ed.), *Advances in Written Text Analysis*. London: Routledge

Sinclair, J. 2004. The lexical item. In J. Sinclair and R. Carter (eds.), *Trust the Text*. London: Routledge

Sinclair, J. *et al.* (eds.) 1990. *Collins Cobuild English Grammar*. London: HarperCollins

Sinclair, J. McH. and Mauranen, A. 2006. *Linear Unit Grammar*. Amsterdam: John Benjamins

Smith, F. 1985. *Reading*. Cambridge: Cambridge University Press

Sperber, D., Caro, F. and Girotto, V. 1995. Relevance theory explains the selection task. *Cognition 57* p.31-95

Sperber, D. and Girotto, V. 2002. Use or misuse of the selection task? Rejoinder to Fiddick, Cosmides, and Tooby. *Cognition 85* p.277-290

Steedman, M. 1991. Structure and intonation. *Language 67/2* p.260-296

Steedman, M. 2000. Information structure and the syntax-morphology interface. *Linguistic Inquiry 31/4* p.649-689

Steels, L. 1998. The origins and ontologies and communication conventions in multi-agent systems. *Autonomous Agents and Multi-Agent Systems 1* p.169-194

Steels, L. 2000. Language as a complex adaptive system. In M. Shoenhauer *et al.* (eds.), *Parallel Problem Solving from Nature – Lecture Notes in Computer Science Vol. 1917*. Berlin: Springer-Verlag

Stotsky, S. 1983. Types of lexical cohesion in expository writing: Implications for developing the vocabulary of academic discourse. *College Composition and Communication 34/4* p.430-446

Strube, M. 1998. Never look back: An alternative to centering. *Proceedings of COLING '98/ACL'98*, Montreal, Canada. p.1251-1257. Retrieved 20 November 2003 from http://citeseer.nj.nec.com/cache/papers/cs/27840/http:zSzzSzacl.ldc.upenn.eduzSzPzSzP98zSzP98-2204.pdf/strube98never.pdf

Strube, M. and Hahn, U. 1999. Functional centering - grounding referential coherence in information structure. *Computational Linguistics 25/3* p.309-344

Svoboda, A. 1981. *Diatheme*. Brno: Brno University

Swales, J. 1990. *Genre Analysis*. Cambridge: Cambridge University Press

Taboada, M. 2004. *Building Coherence and Cohesion*. Amsterdam: John Benjamins

Tadros, A. 1985. *Prediction in Text*. Birmingham: English Language Research

Tadros, A. 1989. Predictive categories in university textbooks. *English for Specific Purposes 8* p.17-31

Tannen, D. 1979. What's in a frame? Surface evidence for underlying expectations. In R. Freedle (ed.). *New Directions in Discourse Processing*. New Jersey: Ablex

Tannenbaum, A.S. 1995. *Distributed Operating Systems*. Upper Saddle River: Prentice Hall

Taylor Torsello, C. 1996. On the logical metafunction. *Functions of Language 3/2* p.151-183

Tetreault, J. 2001. A corpus-based evaluation of centering and pronoun resolution. *Computational Linguistics 27/4* p.507-20

Thibault, P.J. 1999. Communicating and interpreting relevance through discourse negotiation: An alternative to relevance theory. *Journal of Pragmatics 31* p.557-594

Thibault, P.J. 2004a. *Brain, Mind and the Signifying Body – An Ecosocial Semiotic Theory*. London: Continuum

Thibault, P.J. 2004b. *Agency and Consciousness in Discourse – Self-Other Dynamics as a Complex System*. London: Continuum

Thibault, P.J. 2005. Brains, bodies, contextualising activity and language: Do humans (and bonobos) have a language faculty and can they do without one? *Linguistics and the Human Sciences 1/1* p.99-125

Thibault, P.J. and van Leeuwen, T. 1996. Grammar, society and the speech act: Renewing the connections. *Journal of Pragmatics 25* p.561-585

Thompson, G. 2006. Theme in text. In K. Brown (ed.), *The Encyclopedia of Language and Linguistics – Second Edition*. Oxford: Elsevier

Thompson, G. 2007. Unfolding Theme: The development of clausal and textual perspectives on Theme. In J. Webster, C.M.I.M. Matthiessen and R. Hasan (eds.), *Continuing Discourse on Language. Vol. 2*. London: Equinox

Thompson, S.A. 1985. Grammar and written discourse: Initial vs. final purpose clauses in English. *Text 5/1-2* p.55-83

Tomasino, B., Werner, C. J., Weiss, P. H. and Finka, G. R. 2007. Stimulus properties matter more than perspective: An fMRI study of mental imagery and silent reading of action phrases. *NeuroImage 36* p.T128-T141

Truss, L. 2003. *Eats, Shoots and Leaves*. London: Profile Books

Tuttle, S.G. 2008. Phonetics and word definition in Ahtna Athabascan. *Linguistics 46/2* p.439-470

Underwood, G. and Holt, P. 1979. Cognitive skills in the reading process: A review. *Journal of Research in Reading 2/2* p.82-94

Vachek, J. 1987. Written language seen from the functionalist angle. In R. Dirven and V. Fried (eds.), *Functionalism in Linguistics*. Amsterdam: John Benjamins

Vachek, J. 1989. *Written Language Revisited*. Amsterdam: John Benjamins

Varela, F.J., Thompson, E. and Rosch, E. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, Mass: The MIT Press

Vasconcellos, M. 1992. The theme as message onset: Its structure and characteristics. *Linguistics 30* p.147-163

Vallduvi, E. 1993. Information Packaging: A survey. Report prepared for WOPIS Retrieved on 2 October 2007 from http://www.hcrc.ed.ac.uk/publications/rp-44.ps.gz

Vallduvi, E. and Engdahl, E. 1996. The linguistic realization of information packaging. *Linguistics 34* p.459-519

Vernon, D. and Furlong, D. 2007. Philosophical foundations of AI. In M. Lungarella *et al.* (eds.) *50 Years of AI*. Berlin: Springer-Verlag

Vigneau, M. *et al.* 2006. Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage 30* p.1414-1432

Westermann, G., Ruh, N. & Plinkett, K. 2009. Connectionist approaches to language learning. *Linguistics 47/2* p.413-452

Whittaker, R. 1995. Theme, processes and the realization of meanings in academic articles. In M. Ghadessy (ed.), *Thematic Development in English Texts*. London: Pinter

Wilson, D. and Matsui, T. 1998. Recent approaches to bridging: Truth, coherence, relevance. *UCL Working Papers in Linguistics 10* p.1-28

Wilson, D., and Sperber, D. 1986. *Pragmatics: An Overview. Centre for Language and Communication Studies Occasional Paper no.16*. Dublin: CLCS, Trinity College

Winter, E. 1977. A clause relational approach to English texts: a study of some predictive lexical items in written discourse. *Instructional Science 6* p.1-92

Winter, E. 1992. The notion of unspecific versus specific as one way of analysing the information of a fund-raising letter. In W. Mann and S.A. Thompson (eds.), *Discourse Description*. Amsterdam: John Benjamins

Witte, S. and Faigley, L. 1981. Coherence, cohesion and writing quality. *College Composition and Communication 32/2* p.189-204

Appendices

Description & Guide to CD-Rom

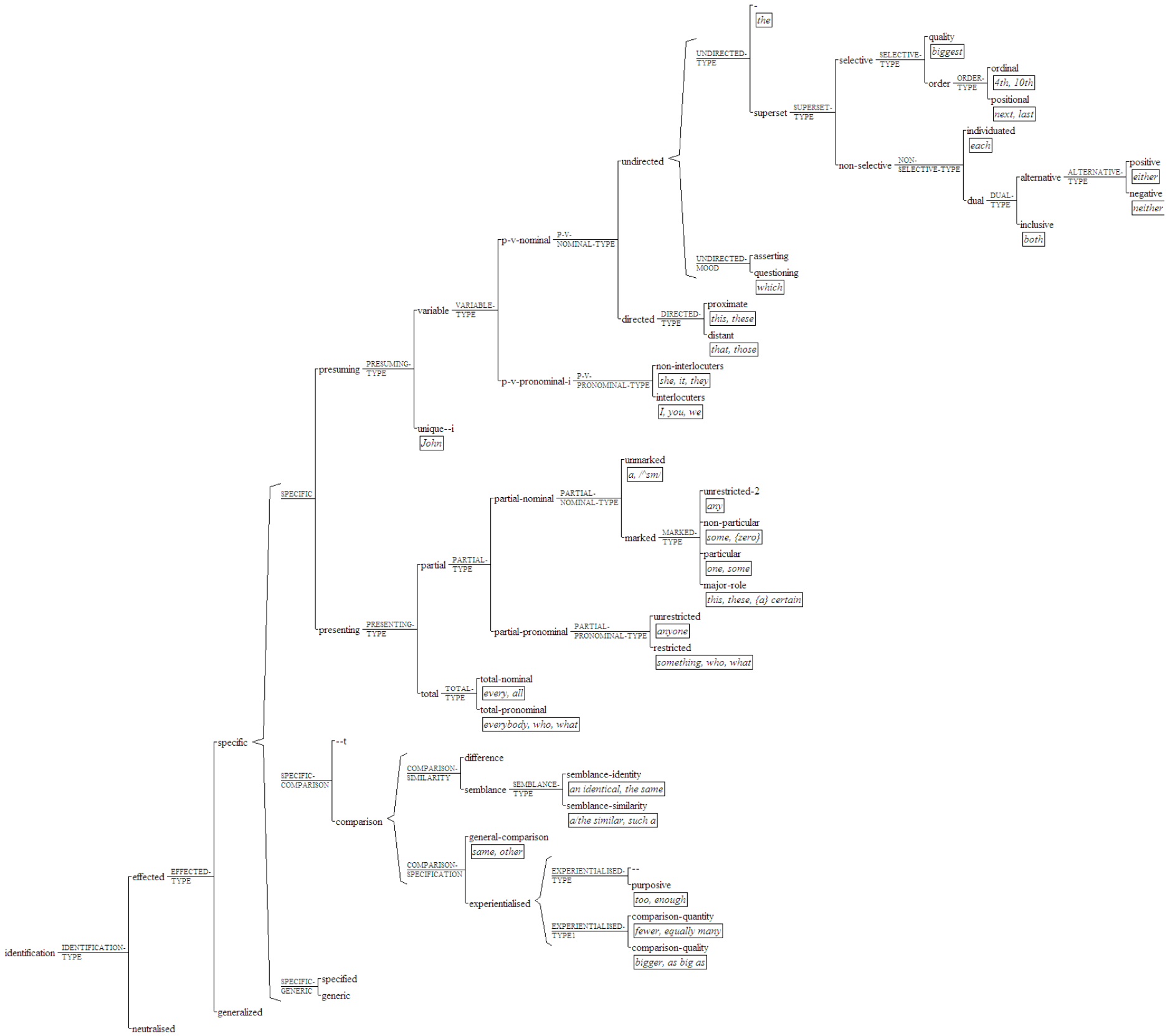
All appendices to this study can be found on the CD-Rom inserted on the inside back cover. This section offers a description of the contents and instructions for use.

Insert the CD-Rom into a computer. If the home page does not automatically open in your default web browser, open the file on the CD-Rom labelled "index.htm". The home page provides links to a pdf version and a Word version of the whole thesis. It also contains links to all appendices.

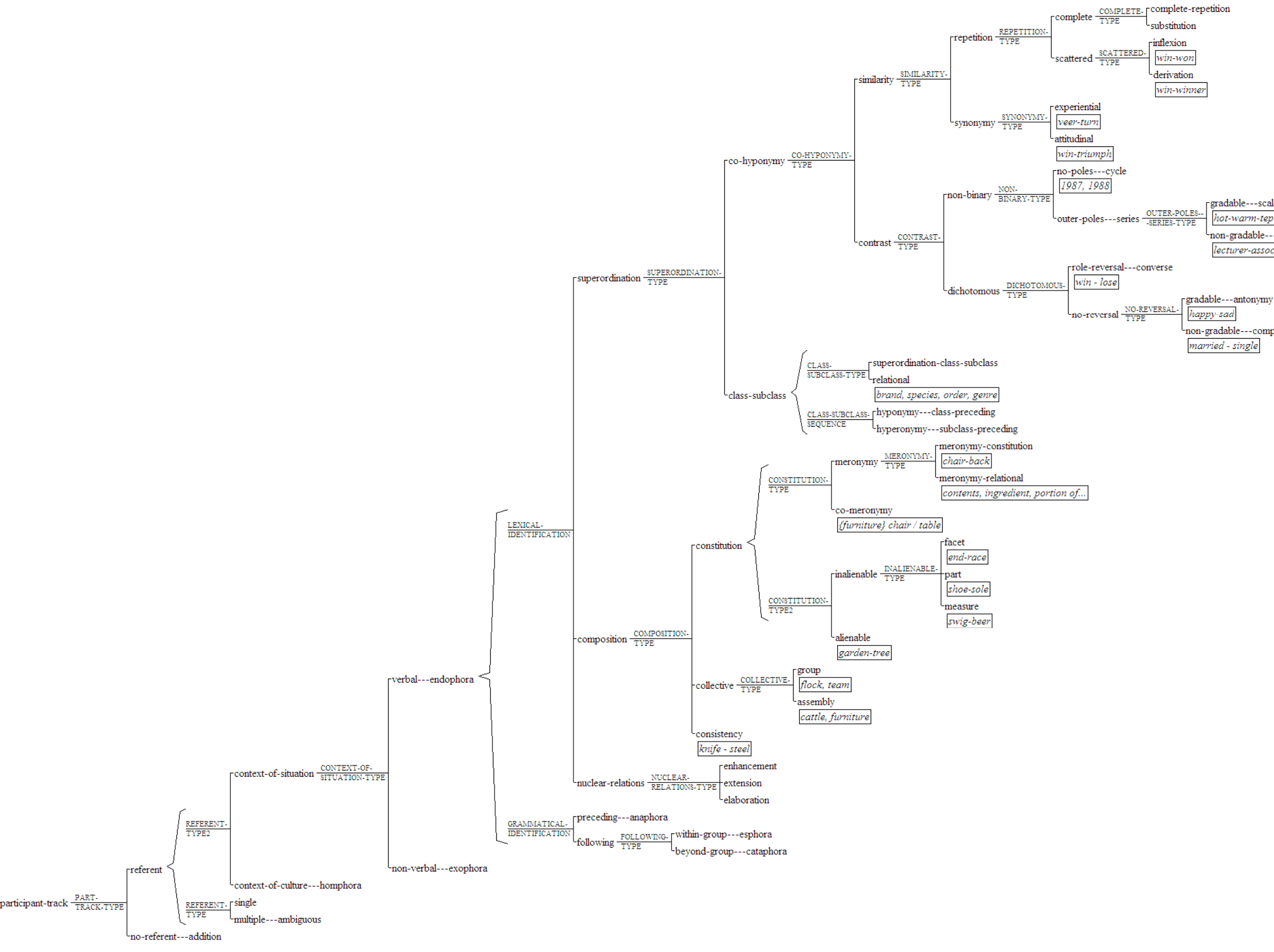
The CD-Rom contains all appendices referred to in this thesis. They are organised by chapter. Each chapter has its own 'home page', or index, containing links to appendices both in html versions and versions that can be easily printed (MS Word). Each chapter index has links to the CD-Rom 'home page', as well as links to the previous and following chapter lists. Similarly, at the end of each appendix there are links to the list of appendices for the chapter, and to the previous and following appendices. In addition, the Back and Forward buttons on the browser can also be used for navigation.

Finally, the CD-Rom contains the software package used throughout this study, UAM CorpusTools 2.4.2. Appendix 5 provides instructions on how to install the software and access all of the textual data used in the study, including texts, analytical schemes and statistics. It also provides a link to download the latest version of the software.

Appendix 2.1 Identification Network



Appendix 2.2 Tracking Network



Appendix 2.3 Participant Analysis of 'Amplifier Noise' text

#	Text	Tracking	Identification
1	AMPLIFER NOISE	no-referent---addition	neutralised
2	In almost every area of	single within-group---esphora hyperonymy---subclass-preceding superordination-class-subclass	generic general-comparison difference non-particular
3	measurement	no-referent---addition	generic --t unmarked
4	the ultimate limit of	multiple---ambiguous within-group---esphora measure meronymy-relational	specified --t asserting quality
5	detectability of	single within-group---esphora hyperonymy---subclass-preceding superordination-class-subclass	generic --t unmarked
6	weak signals	no-referent---addition	generic --t unmarked
7	is set by	no tracking	no identification
8	noise -	no-referent---addition	generic --t unmarked
9	unwanted signals	no-referent---addition	specified --t non-particular
10	that obscure	no tracking	no identification
11	the desired signal.	single preceding---anaphora role-reversal---converse	--t specified - asserting
12	Even if	no tracking	no identification
13	the quantity being measured	single within-group---esphora alienable meronymy-relational	--t specified - asserting
14	is not	no tracking	no identification
15	weak,	no tracking	no identification
16	the presence of	single within-group---esphora elaboration	specified --t - asserting
17	noise	single preceding---anaphora complete-repetition	generic --t - asserting
18	degrades	no tracking	no identification
19	the accuracy of	single within-group---esphora hyperonymy---subclass-preceding relational	specified --t - asserting
20	the measurement	single preceding---anaphora complete-repetition	--t specified - asserting
21	. Some forms of	single within-group---esphora hyperonymy---subclass-preceding superordination-class-subclass	--t specified particular
22	noise	single preceding---anaphora complete-repetition	generic --t - asserting
23	are	no tracking	no identification
24	unavoidable	no tracking	no identification
25	(e.g.	no tracking	no identification
26	real fluctuations in	no-referent---addition	--t specified unmarked
27	the quantity being measured	single preceding---anaphora hyperonymy---subclass-preceding superordination-class-subclass	--t specified - asserting
28	), and	no tracking	no identification
29	they	single preceding---anaphora substitution	--t specified non-interlocuters
30	can be overcome only with	no tracking	no identification
31	the techniques of	single within-group---esphora hyponymy---class-preceding superordination-class-subclass	--t specified - asserting
32	signal averaging	no-referent---addition	--t specified unmarked
33	and	no tracking	no identification
34	bandwidth narrowing	no-referent---addition	--t specified unmarked

#	Text	Tracking	Identification
35	, which	no tracking	no identification
36	we	single non-verbal---exophora	--t specified interlocuters
37	will discuss	no tracking	no identification
38	in chapter 15.	non-verbal---exophora single	specified --t unmarked
39	Other forms of	single within-group---esphora hyperonymy---subclass-preceding relational	specified general-comparison difference asserting -
40	noise	single preceding---anaphora complete-repetition	generic --t - asserting
41	(e.g.	no tracking	no identification
42	radiofrequency interference and	no-referent---addition	generic --t unmarked
43	"ground loops"	no-referent---addition	--t specified unique--i
44	) can be reduced	no tracking	no identification
45	or eliminated by	no tracking	no identification
46	a variety of	no-referent---addition	--t specified unmarked
47	tricks,	no-referent---addition	--t specified unmarked
48	including	no tracking	no identification
49	filtering	no-referent---addition	--t generic unmarked
50	and	no tracking	no identification
51	careful attention to	no-referent---addition	--t specified unmarked
52	wiring configuration	no-referent---addition	--t specified unmarked
53	and parts location	no-referent---addition	--t specified unmarked
54	. Finally,	no tracking	no identification
55	there	single beyond-group---cataphora elaboration	generalized
56	is	no tracking	no identification
57	noise	single preceding---anaphora complete-repetition	specified --t asserting -
58	that arises	no tracking	no identification
59	in the amplification	multiple---ambiguous within-group---esphora derivation	--t specified - asserting
60	itself,	single preceding---anaphora complete-repetition	--t specified non-interlocuters
61	and	no tracking	no identification
62	it	single preceding---anaphora substitution	--t specified non-interlocuters
63	can be reduced	no tracking	no identification
64	the techniques of	single within-group---esphora hyponymy---class-preceding superordination-class-subclass	--t specified - asserting
65	low-noise amplifier design.	no-referent---addition	--t specified unmarked
66	Although	no tracking	no identification
67	the techniques of	single within-group---esphora hyponymy---class-preceding superordination-class-subclass	--t specified - asserting
68	signal averaging	single preceding---anaphora complete-repetition	--t specified unmarked
69	can often be used	no tracking	no identification
70	to rescue	no tracking	no identification
71	a signal	no-referent---addition	--t specified unmarked
72	buried in	no tracking	no identification
73	noise	single preceding---anaphora complete-repetition	generic --t - asserting
74	, it	no-referent---addition	generalized
75	pays to	no tracking	no identification
76	begin with	no tracking	no identification

#	Text	Tracking	Identification
77	a system	no-referent---addition	--t specified unmarked
78	that is	no tracking	no identification
79	free of	no tracking	no identification
80	preventable interference	no-referent---addition	--t specified unmarked
81	and that possesses	no tracking	no identification
82	the lowest amplifier noise predictable.	single preceding---anaphora inflexion	specified --t asserting quality
83	We	single non-verbal---exophora	--t interlocuters specified
84	will begin	no tracking	no identification
85	by talking about	no tracking	no identification
86	the origins and	single within-group---esphora hyponymy---class-preceding superordination-class-subclass	generic --t - asserting
87	characteristics of	single within-group---esphora hyponymy---class-preceding superordination-class-subclass	generic --t - asserting
88	the different kinds of	single within-group---esphora hyperonymy---subclass-preceding relational	specified difference general-comparison - asserting
89	noise	multiple---ambiguous preceding---anaphora experiential	--t specified unmarked
90	that afflict	no tracking	no identification
91	electronic circuits.	no-referent---addition	generic --t asserting -
92	Then	no tracking	no identification
93	we	single non-verbal---exophora	--t specified interlocuters
94	will launch into	no tracking	no identification
95	a discussion of	no-referent---addition	--t specified unmarked
96	transistor	no-referent---addition	--t specified - asserting
97	and FET noise,	no-referent---addition	--t specified unmarked
98	including	no tracking	no identification
99	methods for	single within-group---esphora hyponymy---class-preceding superordination-class-subclass	--t specified - asserting
100	low-noise design with	no-referent---addition	--t specified unmarked
101	a given signal source,	no-referent---addition	--t specified unmarked
102	and will present	no tracking	no identification
103	some design examples.	single preceding---anaphora experiential	--t specified particular
104	After	no tracking	no identification
105	a short discussion of	no-referent---addition	--t specified unmarked
106	noise in	single preceding---anaphora complete-repetition	specified --t - asserting
107	differential [amplifiers]	no-referent---addition	--t specified unmarked
108	and feedback amplifiers,	no-referent---addition	--t specified unmarked
109	we	single non-verbal---exophora	--t specified interlocuters
110	will conclude with	no tracking	no identification
111	a section on	no-referent---addition	--t specified unmarked
112	proper grounding and	no-referent---addition	--t specified unmarked
113	shielding and	no-referent---addition	--t specified unmarked
114	the elimination of	single within-group---esphora elaboration	--t specified - asserting
115	interference and	no-referent---addition	generic --t unmarked
116	pickup.	no-referent---addition	generic --t unmarked
117	7.11	no tracking	no identification
118	Origins and	no-referent---addition	neutralised
119	kinds of	no-referent---addition	neutralised

#	Text	Tracking	Identification
120	noise	single preceding---anaphora complete-repetition	neutralised
121	Since	no tracking	no identification
122	the term noise	single preceding---anaphora experiential	--t specified - asserting
123	can be applied to	no tracking	no identification
124	anything that obscures	single within-group---esphora hyponymy---class-preceding superordination-class-subclass	generic --t unrestricted
125	a desired signal,	no-referent---addition	--t specified unmarked
126	noise	single preceding---anaphora complete-repetition	--t specified - asserting
127	can	no tracking	no identification
128	itself	single preceding---anaphora substitution	--t specified non-interlocuters
129	be	no tracking	no identification
130	another signal	single preceding---anaphora complete-repetition	specified general-comparison difference unmarked
131	("interference")	no-referent---addition	--t specified unmarked
132	; most often,	no tracking	no identification
133	however,	no tracking	no identification
134	we	single non-verbal---exophora	--t specified interlocuters
135	use	no tracking	no identification
136	the term	single preceding---anaphora substitution	--t specified - asserting
137	to describe	no tracking	no identification
138	"random" noise of	no-referent---addition	--t specified unmarked
139	a physical (often thermal) origin	no-referent---addition	--t specified unmarked
140	. Noise	single preceding---anaphora complete-repetition	generic --t - asserting
141	can be characterized by	no tracking	no identification
142	its frequency spectrum,	single preceding---anaphora facet meronymy-relational	--t specified unmarked
143	its amplitude distribution,	single preceding---anaphora facet meronymy-relational	--t specified unmarked
144	and the physical mechanism responsible for	single within-group---esphora elaboration	specified --t - asserting
145	its generation.	single preceding---anaphora hyponymy---class-preceding superordination-class-subclass	specified --t unmarked
146	(Acknowledgment: P. Horowitz & W. Hill. 1989. The Art of Electronics. Cambridge: CUP)	no tracking	no identification

Appendix 2.4 Participant Analysis of 'Milling Machine' text

#	Text	Tracking	Identification
1	11.1 Milling-machine elements	no-referent---addition	neutralised
2	The main elements of	single within-group---esphora part meronymy-constitution	generic --t asserting individuated
3	a typical knee-and-column horizontal milling machine	no-referent---addition	specified --t unmarked
4	are shown in	no tracking	no identification
5	Fig. 11.3.	no-referent---addition	specified --t unique--i
6	The elements of	single within-group---esphora part meronymy-constitution	specified --t asserting -
7	a vertical machine	no-referent---addition	specified --t unmarked
8	are	no tracking	no identification
9	the same	single preceding---anaphora substitution	specified general-comparison semblance-identity asserting -
10	except that	no tracking	no identification
11	the spindle head	single non-verbal---exophora	generic --t asserting -
12	is mounted at	no tracking	no identification
13	the top	single within-group---esphora facet meronymy-constitution	specified --t - asserting
14	of the column,	single non-verbal---exophora	--t specified - asserting
15	as shown in	no tracking	no identification
16	Fig. 11.4.	no-referent---addition	specified --t non-particular
17	Column	no-referent---addition	neutralised
18	and base	no-referent---addition	neutralised
19	The column and base	single non-verbal---exophora	specified --t asserting -
20	form	no tracking	no identification
21	the foundation of	single within-group---esphora meronymy-constitution part	--t specified - asserting
22	the complete machine.	single preceding---anaphora experiential	--t specified - asserting
23	Both	single preceding---anaphora substitution	--t specified asserting inclusive
24	are made from	no tracking	no identification
25	cast iron,	no-referent---addition	--t generic unmarked
26	designed with	no tracking	no identification
27	thick sections	no-referent---addition	--t generic unmarked
28	to ensure	no tracking	no identification
29	complete rigidity and	no-referent---addition	generic --t unmarked
30	freedom from	no-referent---addition	--t generic unmarked
31	vibration.	no-referent---addition	--t generic unmarked
32	The base,	single preceding---anaphora complete-repetition	--t specified - asserting
33	upon which	no tracking	no identification
34	the column	single preceding---anaphora complete-repetition	--t specified - asserting
35	is mounted,	no tracking	no identification
36	is also	no tracking	no identification
37	the cutting-fluid reservoir	single non-verbal---exophora	--t specified - asserting
38	and contains	no tracking	no identification
39	the pump	single non-verbal---exophora	--t specified - asserting
40	to circulate	no tracking	no identification
41	the fluid to	single preceding---anaphora facet meronymy-constitution	--t specified - asserting
42	the cutting area.	single non-verbal---exophora	--t specified - asserting

#	Text	Tracking	Identification
43	The column	single preceding---anaphora complete-repetition	--t specified - asserting
44	contains	no tracking	no identification
45	the spindle,	single non-verbal---exophora	--t specified - asserting
46	accurately located in	no tracking	no identification
47	precision bearings.	no-referent---addition	--t generic unmarked
48	The spindle	single preceding---anaphora complete-repetition	--t specified - asserting
49	is driven through	no tracking	no identification
50	a gearbox from	no-referent---addition	--t specified unmarked
51	a vee-belt drive from	no-referent---addition	--t specified unmarked
52	the electric motor	single context-of-culture---homphora	--t specified - asserting
53	housed at	no tracking	no identification
54	the base of	single within-group---esphora part meronymy-constitution	--t specified - asserting
55	the column.	single preceding---anaphora complete-repetition	--t specified - asserting
56	The gearbox	single preceding---anaphora complete-repetition	--t specified - asserting
57	enables	no tracking	no identification
58	a range of	single within-group---esphora measure meronymy-constitution	generic --t unmarked
59	spindle speeds	no-referent---addition	--t generic unmarked
60	to be selected.	no tracking	no identification
61	In the model	single non-verbal---exophora	--t specified - asserting
62	shown,	no tracking	no identification
63	twelve spindle speeds from	single preceding---anaphora complete-repetition	--t specified particular
64	32 to	single preceding---anaphora no-poles---cycle	--t specified unmarked
65	1400rev/min	single preceding---anaphora no-poles---cycle	--t specified unmarked
66	are available.	no tracking	no identification
67	The front of	single within-group---esphora part meronymy-constitution	--t specified - asserting
68	the column	single preceding---anaphora complete-repetition	--t specified - asserting
69	carries	no tracking	no identification
70	the guideways	no-referent---addition	--t specified - asserting
71	upon which	no tracking	no identification
72	the knee	single non-verbal---exophora	--t specified - asserting
73	is located and guided in	no tracking	no identification
74	a vertical direction.	no-referent---addition	--t specified unmarked
75	Figure 11.3	no-referent---addition	neutralised
76	Main machine elements of horizontal milling machine	no-referent---addition	neutralised
77	Figure 11.4	no-referent---addition	neutralised
78	Top of column of vertical milling machine	no-referent---addition	neutralised
79	Knee	single preceding---anaphora complete-repetition	neutralised
80	The knee,	single preceding---anaphora complete-repetition	specified --t asserting -
81	mounted on	no tracking	no identification
82	the column guideways,	single preceding---anaphora part meronymy-constitution	--t specified - asserting

#	Text	Tracking	Identification
83	provides	no tracking	no identification
84	the vertical movement of	single within-group---esphora experiential	--t specified - asserting
85	the table.	single non-verbal---exophora	--t specified - asserting
86	Power feed	no-referent---addition	--t generic unmarked
87	is available, through	no tracking	no identification
88	a gearbox	no-referent---addition	--t specified unmarked
89	mounted on	no tracking	no identification
90	the side, from	single preceding---anaphora meronymy-constitution part	--t specified - asserting
91	a separate built-in motor,	no-referent---addition	specified difference general-comparison unmarked
92	providing	no tracking	no identification
93	a range of	single within-group---esphora measure meronymy-constitution	specified --t unmarked
94	twelve feed rates from	no-referent---addition	--t specified unmarked
95	6 to	single preceding---anaphora no-poles---cycle	--t specified unmarked
96	250 mm/min.	single preceding---anaphora no-poles---cycle	--t specified unmarked
97	Drive	no-referent---addition	generic --t unmarked
98	is through	no tracking	no identification
99	a leadscrew,	no-referent---addition	--t specified unmarked
100	whose	no tracking	no identification
101	bottom end	single preceding---anaphora meronymy-constitution part	--t specified unmarked
102	is fixed to	no tracking	no identification
103	the machine base.	single preceding---anaphora derivation	--t specified - asserting
104	Provision	no-referent---addition	--t generic unmarked
105	is made to raise and lower	no tracking	no identification
106	the knee	single preceding---anaphora complete-repetition	--t specified - asserting
107	by hand	no-referent---addition	--t generic unmarked
108	through	no tracking	no identification
109	a leadscrew and nut	no-referent---addition	--t specified unmarked
110	operated by	no tracking	no identification
111	a handwheel at	no-referent---addition	--t specified unmarked
112	the front.	single preceding---anaphora meronymy-constitution part	--t specified - asserting
113	The knee	single preceding---anaphora complete-repetition	--t specified - asserting
114	has	no tracking	no identification
115	guideways on	no-referent---addition	generic --t unmarked
116	its top surface	single preceding---anaphora meronymy-constitution part	--t specified non-interlocuters
117	giving	no tracking	no identification
118	full-width support to	no-referent---addition	--t generic unmarked
119	the saddle	single non-verbal---exophora	--t specified - asserting
120	and guiding	no tracking	no identification
121	it in	multiple---ambiguous preceding---anaphora substitution	--t specified non-interlocuters
122	a transverse direction.	no-referent---addition	specified --t unmarked
123	A lock	no-referent---addition	--t specified unmarked
124	is provided to clamp	no tracking	no identification

#	Text	Tracking	Identification
125	the knee in	single preceding---anaphora complete-repetition	--t specified - asserting
125	any vertical position on	single within-group---esphora attitudinal	--t specified unrestricted-2
126	the column.	single preceding---anaphora complete-repetition	--t specified - asserting
127	Saddle	single non-verbal---exophora	neutralised
128	The saddle,	single preceding---anaphora complete-repetition	specified --t asserting -
129	mounted on	no tracking	no identification
130	the knee guideways,	single preceding---anaphora derivation	--t specified - asserting
131	provides	no tracking	no identification
132	the transverse movement of	single within-group---esphora role-reversal---converse	--t specified - asserting
133	the table.	single preceding---anaphora complete-repetition	--t specified - asserting
134	Power feed	no-referent---addition	--t generic unmarked
135	is provided through	no tracking	no identification
136	the gearbox on	single within-group---esphora meronymy-constitution part	--t specified - asserting
137	the knee.	single preceding---anaphora complete-repetition	--t specified - asserting
138	A range of	single within-group---esphora measure meronymy-constitution	--t specified unmarked
139	twelve feeds	no-referent---addition	--t specified unmarked
140	is available, from	no tracking	no identification
141	12 to	single preceding---anaphora no-poles---cycle	--t specified unmarked
142	500 mm/min.	single preceding---anaphora no-poles---cycle	--t specified unmarked
143	Alternative hand movement	single preceding---anaphora non-gradable---complementarity	specified difference general-comparison unmarked
144	is provided through	no tracking	no identification
145	a leadscrew and nut by	no-referent---addition	--t specified unmarked
146	a handwheel at	no-referent---addition	--t specified unmarked
147	the front of	single within-group---esphora meronymy-constitution part	--t specified - asserting
148	the knee.	single preceding---anaphora complete-repetition	--t specified - asserting
149	Clamping of	single within-group---esphora enhancement	--t generic unmarked
150	the saddle to	single preceding---anaphora complete-repetition	--t specified - asserting
151	the knee	single preceding---anaphora complete-repetition	--t specified - asserting
152	is achieved by	no tracking	no identification
153	two clamps on	single preceding---anaphora inflexion	--t specified unmarked
154	the side of	single preceding---anaphora meronymy-constitution part	--t specified - asserting
155	the saddle.	single preceding---anaphora complete-repetition	--t specified - asserting
156	The saddle	single preceding---anaphora complete-repetition	--t specified - asserting
157	has	no tracking	no identification
158	dovetail guideways on	no-referent---addition	--t specified unmarked

#	Text	Tracking	Identification
159	its upper surface,	single preceding---anaphora meronymy-constitution part	--t specified non-interlocuters
160	at right angles to	no tracking	no identification
161	the knee guideways,	single preceding---anaphora derivation	--t specified - asserting
162	to provide	no tracking	no identification
163	a guide to	no-referent---addition	--t generic unmarked
164	the table in	single preceding---anaphora complete-repetition	--t specified - asserting
165	a longitudinal direction.	multiple---ambiguous preceding---anaphora experiential	--t specified unmarked
166	Table	single preceding---anaphora complete-repetition	neutralised
167	The table	single preceding---anaphora complete-repetition	specified --t asserting -
168	provides	no tracking	no identification
169	the surface	no-referent---addition	specified --t asserting -
170	upon which	no tracking	no identification
171	all workpieces and workholding equipment	no-referent---addition	specified --t total-nominal
172	are located and clamped.	no tracking	no identification
173	A series of	single within-group---esphora meronymy-constitution measure	--t specified unmarked
174	tee slots	no-referent---addition	--t specified unmarked
175	is provided for	no tracking	no identification
176	this purpose.	single preceding---anaphora attitudinal	--t specified - asserting
177	The dovetail guides on	single within-group---esphora meronymy-constitution part	--t specified - asserting
178	the undersurface	single preceding---anaphora meronymy-constitution part	--t specified - asserting
179	locate in	no tracking	no identification
180	the guideways on	single within-group---esphora part meronymy-constitution	--t specified - asserting
181	the saddle,	single preceding---anaphora complete-repetition	--t specified - asserting
182	giving	no tracking	no identification
183	straight-line movement to	multiple---ambiguous preceding---anaphora experiential	--t specified unmarked
184	the table in	single preceding---anaphora complete-repetition	--t specified - asserting
185	a longitudinal direction at	no-referent---addition	--t specified unmarked
186	right angles to	no tracking	no identification
187	the saddle movement.	single preceding---anaphora experiential	--t specified - asserting
188	Power feed	no-referent---addition	--t generic unmarked
189	is provided from	no tracking	no identification
190	the knee gearbox,	single preceding---anaphora experiential	--t specified - asserting
191	through	no tracking	no identification
192	the saddle,	single preceding---anaphora complete-repetition	--t specified - asserting
193	to	no tracking	no identification
194	the table leadscrew.	single preceding---anaphora meronymy-constitution part	--t specified - asserting
195	Alternative hand feed	single preceding---anaphora non-gradable---complementarity	--t specified unmarked
196	is provided by	no tracking	no identification

#	Text	Tracking	Identification
197	a handwheel at	no-referent---addition	--t specified unmarked
198	each end of	single within-group---esphora meronymy-constitution part	--t specified asserting individuated
199	the table.	single preceding---anaphora complete-repetition	--t specified - asserting
200	Stops at	no-referent---addition	--t generic unmarked
201	the front of	single preceding---anaphora meronymy-constitution part	--t specified - asserting
202	the table	single preceding---anaphora complete-repetition	--t specified - asserting
203	can be set to disengage	no tracking	no identification
204	the longitudinal feed	single preceding---anaphora experiential	--t specified - asserting
205	automatically in	no tracking	no identification
206	each direction.	single preceding---anaphora role-reversal---converse	--t specified asserting individuated
207	Spindle	no-referent---addition	neutralised
208	The spindle,	single preceding---anaphora complete-repetition	specified --t asserting -
209	accurately mounted in	no tracking	no identification
210	precision bearings,	no-referent---addition	--t specified unmarked
211	provides	no tracking	no identification
212	the drive for	single within-group---esphora facet meronymy-relational	--t specified - asserting
213	the milling cutters.	no-referent---addition	--t specified - asserting
214	Cutters	single preceding---anaphora complete-repetition	--t generic - asserting
215	can be mounted straight on	no tracking	no identification
216	the spindle nose	single preceding---anaphora meronymy-constitution part	--t specified - asserting
217	or in	no tracking	no identification
218	cutterholding devices	no-referent---addition	--t generic - asserting
219	which in turn are mounted in	no tracking	no identification
220	the spindle,	single preceding---anaphora complete-repetition	--t specified - asserting
221	held in	no tracking	no identification
222	position by	no tracking	no identification
223	a drawbolt	no-referent---addition	--t specified unmarked
224	passing through	no tracking	no identification
225	the hollow spindle.	single preceding---anaphora substitution	--t specified - asserting
226	Spindles of	single preceding---anaphora inflexion	--t generic - asserting
227	milling machines	single preceding---anaphora inflexion	--t generic - asserting
228	have	no tracking	no identification
229	a standard spindle nose,	single preceding---anaphora substitution	--t specified unmarked
230	shown in	no tracking	no identification
231	Fig. 11.5,	single non-verbal---exophora	--t specified unmarked
232	to allow for	no tracking	no identification
233	easy interchange of	no-referent---addition	--t specified unmarked
234	cutters and	single preceding---anaphora complete-repetition	--t generic - asserting
235	cutter-holding devices.	single preceding---anaphora complete-repetition	--t generic - asserting

#	Text	Tracking	Identification
236	The bore of	single within-group---esphora facet meronymy-constitution	--t asserting - generic
237	the nose	single preceding---anaphora complete-repetition	--t specified - asserting
238	is tapered to provide	no tracking	no identification
239	accurate location,	no-referent---addition	--t generic unmarked
240	the angle of	single within-group---esphora meronymy-constitution measure	--t specified - asserting
241	taper	no-referent---addition	--t generic unmarked
242	being	no tracking	no identification
243	16° 36'.	single preceding---anaphora no-poles---cycle	--t specified unmarked
244	The diameter of	single within-group---esphora meronymy-constitution measure	--t specified - asserting
245	the taper	single preceding---anaphora complete-repetition	--t specified - asserting
246	depends on	no tracking	no identification
247	the size of	single within-group---esphora meronymy-constitution measure	--t specified - asserting
248	the machine	single preceding---anaphora substitution	--t specified - asserting
249	and may be	no tracking	no identification
250	30, 40, or 50 IST (International Standard Taper).	single preceding---anaphora no-poles---cycle	--t specified unmarked
251	Due to	no tracking	no identification
252	their steepness of	single beyond-group---cataphora meronymy-constitution facet	--t specified non-interlocuters
253	angle,	single preceding---anaphora complete-repetition	--t generic - asserting
254	these tapers	single preceding---anaphora complete-repetition	--t specified proximate
255	- known as	no tracking	no identification
256	non-stick or	no-referent---addition	--t generic unmarked
257	self-releasing	no-referent---addition	--t generic unmarked
258	- cannot be relied upon to transmit	no tracking	no identification
259	the drive to	single preceding---anaphora complete-repetition	--t specified - asserting
260	the cutter or	single preceding---anaphora complete-repetition	--t specified - asserting
261	cutter-holding device.	single preceding---anaphora complete-repetition	--t specified - asserting
262	Two driving keys	no-referent---addition	--t specified unmarked
263	are provided to transmit	no tracking	no identification
264	the drive.	single preceding---anaphora complete-repetition	--t specified - asserting
265	Figure 11.5	no-referent---addition	neutralised
266	Standard milling-machine spindle nose	no-referent---addition	neutralised
267	Cutters	single preceding---anaphora complete-repetition	specified --t unmarked
268	which are mounted directly on	no tracking	no identification
269	the spindle nose	single preceding---anaphora complete-repetition	--t specified - asserting
270	are located on	no tracking	no identification
271	a centring arbor, and	no-referent---addition	--t specified unmarked

#	Text	Tracking	Identification
272	four tapped holes	no-referent---addition	--t specified unmarked
273	are provided to hold	no tracking	no identification
274	the cutter	single preceding---anaphora complete-repetition	--t specified - asserting
275	in position.	no tracking	no identification
276	The two keys	single preceding---anaphora substitution	--t specified - asserting
277	again provide	no tracking	no identification
278	the means of	single within-group---esphora elaboration	generic --t - asserting
279	transmitting	single preceding---anaphora derivation	--t generic - asserting
280	the drive.	single preceding---anaphora complete-repetition	--t specified - asserting
281	The spindle of	single within-group---esphora meronymy-constitution part	--t specified - asserting
282	a horizontal machine	no-referent---addition	--t specified unmarked
283	is fixed and cannot be adjusted in	no tracking	no identification
284	an axial direction,	single preceding---anaphora experiential	--t specified unmarked
285	i.e.	no tracking	no identification
286	along its axis.	no-referent---addition	--t specified non-interlocuters
287	On vertical machines,	single preceding---anaphora role-reversal---converse	--t specified - asserting
288	provision	no-referent---addition	--t generic unmarked
289	is made for	no tracking	no identification
290	axial movement,	single preceding---anaphora complete-repetition	--t generic - asserting
291	which is controlled by	no tracking	no identification
293	a handwheel on	no-referent---addition	--t specified unmarked
294	the spindle head.	single preceding---anaphora complete-repetition	--t specified - asserting
295	The spindle	single preceding---anaphora complete-repetition	--t specified - asserting
296	runs in	no tracking	no identification
297	a quill	no-referent---addition	--t specified unmarked
298	which is moved through	no tracking	no identification
299	a rack and pinion	no-referent---addition	--t specified unmarked
300	in the same way as	single preceding---anaphora substitution	specified general-comparison semblance-identity - asserting
301	a drilling machine spindle	no-referent---addition	--t specified unmarked
302	(see	no tracking	no identification
303	Fig. 7.2).	no-referent---addition	--t specified unmarked
304	A locking bolt	no-referent---addition	--t specified unmarked
305	is provided to lock	no tracking	no identification
306	the quill in	single preceding---anaphora complete-repetition	--t specified - asserting
307	any position along	no-referent---addition	--t specified unrestricted-2
308	its operating length.	single preceding---anaphora meronymy-constitution facet	--t specified non-interlocuters
309	Overarm and arbor support	single non-verbal---exophora	neutralised
310	The majority of	single within-group---esphora hyponymy---class-preceding superordination-class-subclass	generic --t asserting -
311	cutters	single preceding---anaphora complete-repetition	--t generic - asserting

#	Text	Tracking	Identification
312	used on	no tracking	no identification
313	horizontal machines	single preceding---anaphora complete-repetition	--t generic - asserting
314	are held on	no tracking	no identification
315	an arbor	no-referent---addition	--t specified unmarked
316	which is located and held in	no tracking	no identification
317	the spindle.	single preceding---anaphora complete-repetition	--t specified - asserting
318	Due to	no tracking	no identification
319	the length of	single within-group---esphora meronymy-constitution measure	--t specified - asserting
320	the arbors	single preceding---anaphora complete-repetition	--t specified - asserting
321	used,	no tracking	no identification
322	support	no-referent---addition	--t generic - asserting
323	is required at	no tracking	no identification
324	to prevent	no tracking	no identification
324	the outer end	single preceding---anaphora meronymy-constitution part	specified difference -- comparison-quality - asserting
326	deflection	no-referent---addition	--t generic unmarked
327	when cutting takes place.	no tracking	no identification
328	Support	single preceding---anaphora complete-repetition	--t generic - asserting
329	is provided by	no tracking	no identification
330	an arbor-support bracket,	no-referent---addition	--t specified unmarked
331	clamped to	no tracking	no identification
332	an overarm	no-referent---addition	--t specified unmarked
333	which is mounted on	no tracking	no identification
334	top of	single within-group---esphora meronymy-constitution part	--t specified - asserting
335	the column in	single preceding---anaphora complete-repetition	--t specified - asserting
336	a dovetail slide.	no-referent---addition	--t specified unmarked
337	The overarm	single preceding---anaphora complete-repetition	--t specified - asserting
338	is adjustable in or out for	no tracking	no identification
339	different lengths of	single within-group---esphora meronymy-constitution measure	specified difference -- comparison-quantity unmarked
340	arbor,	single preceding---anaphora complete-repetition	--t generic - asserting
341	or can be fully pushed in when	no tracking	no identification
342	arbor support	single preceding---anaphora derivation	--t generic - asserting
343	is not required.	no tracking	no identification
344	Two clamping bolts	no-referent---addition	--t specified unmarked
345	are provided to lock	no tracking	no identification
346	the overarm in	single preceding---anaphora complete-repetition	--t specified - asserting
347	any position.	no-referent---addition	--t specified unrestricted-2
348	The arbor support	single preceding---anaphora derivation	--t specified - asserting
349	is located in	no tracking	no identification
350	the overarm dovetail	single preceding---anaphora derivation	--t specified - asserting
351	and is locked	no tracking	no identification
352	by means of	no tracking	no identification

#	Text	Tracking	Identification
353	its clamping bolt.	single preceding---anaphora meronymy-constitution part	--t specified non-interlocuters
354	A solid bearing	no-referent---addition	--t specified unmarked
355	is provided in which	no tracking	no identification
356	the arbor	single preceding---anaphora complete-repetition	--t specified - asserting
357	runs during	no tracking	no identification
358	spindle rotation	no-referent---addition	generic --t unmarked

Appendix 2.5 Combinations of Two Systems of Participants for Two Texts (terminal choices)

Features of Participant Tracking	Features of Participant Identification	Count
no tracking	no identification	167
no-referent---addition	--t specified unmarked	64
single preceding---anaphora complete-repetition	--t specified - asserting	52
no-referent---addition	generic --t unmarked	31
single preceding---anaphora complete-repetition	generic --t - asserting	14
no-referent---addition	neutralised	13
single non-verbal---exophora	--t specified - asserting	10
single within-group---esphora meronymy-constitution part	--t specified - asserting	10
single preceding---anaphora meronymy-constitution part	--t specified - asserting	8
single preceding---anaphora no-poles---cycle	--t specified unmarked	8
single non-verbal---exophora	--t specified interlocuters	6
single preceding---anaphora derivation	--t specified - asserting	5
single preceding---anaphora experiential	--t specified - asserting	5
no-referent---addition	--t specified - asserting	4
single preceding---anaphora meronymy-constitution facet	--t specified non-interlocuters	4
single preceding---anaphora substitution	--t specified - asserting	4
single within-group---esphora hyponymy---class-preceding superordination-class-subclass	--t specified - asserting	4
single within-group---esphora meronymy-constitution measure	--t specified - asserting	4
multiple---ambiguous preceding---anaphora experiential	--t specified unmarked	3
no-referent---addition	generic --t - asserting	3
single preceding---anaphora complete-repetition	neutralised	3
single preceding---anaphora substitution	--t specified non-interlocuters	3
single within-group---esphora elaboration	--t specified - asserting	3
single within-group---esphora hyperonymy---subclass-preceding relational	specified general-comparison difference asserting -	3
single within-group---esphora hyponymy---class-preceding superordination-class-subclass	generic --t - asserting	3
single within-group---esphora measure meronymy-constitution	--t specified unmarked	3
no-referent---addition	--t specified non-particular	2
no-referent---addition	--t specified unique--i	2
no-referent---addition	--t specified unrestricted-2	2
single non-verbal---exophora	neutralised	2
single non-verbal---exophora	--t specified unmarked	2
single preceding---anaphora complete-repetition	--t specified unmarked	2
single preceding---anaphora derivation	generic --t - asserting	2
single preceding---anaphora facet meronymy-relational	--t specified unmarked	2
single preceding---anaphora inflexion	generic --t - asserting	2
single preceding---anaphora role-reversal---converse	--t specified - asserting	2
single preceding---anaphora substitution	specified general-comparison semblance-identity - asserting	2
single within-group---esphora meronymy-constitution part	--t specified asserting individuated	2

Features of Participant Tracking	Features of Participant Identification	Count
multiple---ambiguous preceding---anaphora substitution	--t specified non-interlocuters	1
multiple---ambiguous within-group---esphora derivation	--t specified - asserting	1
multiple---ambiguous within-group---esphora measure meronymy-relational	--t specified asserting quality	1
no-referent---addition	generalized	1
no-referent---addition	specified difference general-comparison unmarked	1
no-referent---addition	--t specified non-interlocuters	1
no-referent---addition	--t specified total-nominal	1
single beyond-group---cataphora meronymy-constitution facet	--t specified non-interlocuters	1
single context-of-culture---homphora	--t specified - asserting	1
single non-verbal---exophora	generic --t asserting -	1
single preceding---anaphora attitudinal	--t specified - asserting	1
single preceding---anaphora complete-repetition	--t specified non-interlocuters	1
single preceding---anaphora complete-repetition	--t specified particular	1
single preceding---anaphora complete-repetition	--t specified proximate	1
single preceding---anaphora experiential	--t specified particular	1
single preceding---anaphora experiential	--t specified unmarked	1
single preceding---anaphora facet meronymy-constitution	--t specified - asserting	1
single preceding---anaphora hyperonymy---subclass-preceding superordination-class-subclass	--t specified - asserting	1
single preceding---anaphora hyponymy---class-preceding superordination-class-subclass	--t specified unmarked	1
single preceding---anaphora inflexion	--t specified asserting quality	1
single preceding---anaphora inflexion	--t specified unmarked	1
single preceding---anaphora meronymy-constitution part	specified difference -- comparison-quality - asserting	1
single preceding---anaphora meronymy-constitution part	--t specified unmarked	1
single preceding---anaphora non-gradable---complementarity	specified difference general-comparison unmarked	1
single preceding---anaphora non-gradable---complementarity	--t specified unmarked	1
single preceding---anaphora role-reversal---converse	--t specified asserting individuated	1
single preceding---anaphora substitution	--t specified asserting inclusive	1
single preceding---anaphora substitution	--t specified unmarked	1
single within-group---esphora alienable meronymy-relational	--t specified - asserting	1
single within-group---esphora attitudinal	--t specified unrestricted-3	1
single within-group---esphora elaboration	generic --t - asserting	1
single within-group---esphora enhancement	generic --t unmarked	1
single within-group---esphora experiential	--t specified - asserting	1
single within-group---esphora facet meronymy-constitution	--t asserting - generic	1
single within-group---esphora facet meronymy-constitution	--t specified - asserting	1
single within-group---esphora facet meronymy-relational	--t specified - asserting	1

Features of Participant Tracking	Features of Participant Identification	Count
single within-group---esphora hyperonymy---subclass-preceding relational	--t specified - asserting	1
single within-group---esphora hyperonymy---subclass-preceding superordination-class-subclass	generic general-comparison difference non-particular	1
single within-group---esphora hyperonymy---subclass-preceding superordination-class-subclass	--t specified particular	1
single within-group---esphora hyponymy---class-preceding superordination-class-subclass	generic --t unrestricted	1
single within-group---esphora measure meronymy-constitution	--t generic unmarked	1
single within-group---esphora meronymy-constitution measure	specified difference -- comparison-quantity unmarked	1
single within-group---esphora role-reversal---converse	--t specified - asserting	1

Appendix 3.1 χ^2 Contingency Tables

Note:
The following method has been used to derive the value for χ^2

	1	2			χ^2 table
A	*O= Observed Score*	*O= Observed Score*	Total Row A	A1	**$(E-O)^2/E$**
E = Expected Scores	**E = (Row A * Column 1) / Grand Total**	**E = (Row A * Column2) / Grand Total**		A2	**$(E-O)^2/E$**
B	*O= Observed Score*	*O= Observed Score*	Total Row B	B1	**$(E-O)^2/E$**
E = Expected Scores	**E = (Row B * Column1) / Grand Total**	**E = (Row B * Column2) / Grand Total**		B2	**$(E-O)^2/E$**
	Total Column 1	Total Column 2	Grand Total	χ^2 =	**Sum A1-B2 $\Sigma (E-O)^2/E$**
				phi=	**$\sqrt{\chi^2}$/ Grand Total**

Table 1 Contingency Table comparing Thematic Metafunction with Theme Type

a. With Interpersonal scores

	Simple	Complex			χ^2 table
Textual	*138*	*242*	380	A1	**5.428466**
E =	**168.2187**	**211.7813**		A2	**3.451271**
Ideational	*364*	*382*	746	B1	**4.311851**
E =	**330.2399**	**415.7601**		B2	**2.741357**
Interpersonal	*0*	*8*		C1	**3.541446**
E =	**3.541446**	**4.458554**		C2	**2.812984**
	502	632	1134	χ^2 =	**22.28738**
				phi=	**0.140192**

b. Without Interpersonal scores

	Simple	Complex			χ^2 table
Textual	*138*	*242*	380	A1	**5.824968**
E =	**169.4139**	**210.5861**		A2	**2.967142**
Ideational	*364*	*382*	746	B1	**4.686112**
E =	**332.5861**	**413.4139**		B2	**2.387028**
	502	624	1126	χ^2 =	**15.86525**
				phi=	**0.118701**

Table 2 Contingency Table comparing Thematic Metafunction with Thematic Position

	Initial	Non-Initial			χ^2 table
Textual	*216*	*26*	242	A1	**84.41196**
E =	**116.734**	**125.266**		A2	**53.47564**
Ideational	*85*	*297*	382	B1	**78.66254**
E =	**184.266**	**197.734**		B2	**49.83334**
	301	323	624	χ^2 =	**266.3835**
				phi=	**0.653373**

Table 3 Contingency Table comparing Thematic Metafunction with Participant status

	Participant	Non- Participant			χ^2 table
Textual	*45*	*331*	376	A1	**142.1424**
E =	**223.0643**	**152.9357**		A2	**93.27318**
Ideational	*518*	*55*	573	B1	**207.3217**
E =	**339.9357**	**233.0643**		B2	**136.0435**
	563	386	949	χ^2 =	**578.7807**
				phi=	**0.780951**

Table 4 Contingency Table comparing Participant status with Thematic Position

	Participant	Non- Participant			χ^2 table
Non-initial	*227*	*45*	272	A1	**38.80561**
E =	**150.5626**	**121.4374**		A2	**37.83199**
Initial	*78*	*201*	279	B1	**48.11265**
E =	**154.4374**	**124.5626**		B2	**46.90552**
	305	246	551	χ^2 =	**171.6558**
				phi=	**0.558153**

Table 5 Contingency Table comparing Presenting/Presuming Reference with position in Theme/Rheme– Text MM

	Presenting	Presuming			χ^2 table
Theme	*4*	*25*	29		**5.038095238**
	11.6666667	**17.3333333**			**2.519047619**
Rheme	*31*	*27*	58		**3.391025641**
	23.33333333	**34.66666667**			**1.695512821**
	35	52	87	$\chi 2$ =	**12.64368132**
				phi=	**0.381221288**

Table 6 Contingency Table comparing Presenting/Presuming Reference with position in Theme/Rheme – Text BN

	Presenting	Presuming			χ^2 table
Theme	*14*	*21*	35		**2.190740741**
	20.74074074	**14.25925926**			**1.05035515**
Rheme	*50*	*23*	73		**3.186531987**
	43.25925926	**29.74074074**			**1.527789309**
	64	44	108	critical value=	**7.955417186**
				phi=	**0.271406098**

Table 7 Contingency Table comparing Presenting/Presuming Reference with position in Theme/Rheme – Text Sec

	Presenting	Presuming			χ^2 table
Theme	*25*	*56*	81		**0.8636112**
	30.0983607	**50.9016393**			**0.685808894**
Rheme	*43*	*59*	102		**0.510657058**
	37.90163934	**64.09836066**			**0.405521781**
	68	115	183	critical value=	**2.465598934**
				phi=	**0.116074193**

Table 8 Contingency Table comparing Presenting/Presuming Reference with position in Theme/Rheme – Text AN

	Presenting	Presuming			χ^2 table
Theme	*19*	*56*	75		**0.317871622**
	21.6216216	**53.3783784**			**0.162179399**
Rheme	*45*	*102*	147		**0.128758125**
	42.37837838	**104.6216216**			**0.065692921**
	64	158	222	critical value=	**0.674502067**

Table 9 Contingency Table comparing Presenting/Presuming Reference with position in Theme/Rheme – Text RM

	Presenting	Presuming			χ^2 table
Theme	*61*	*175*	236		**6.39379889**
	84.202946	**151.797054**			**4.023830768**
Rheme	*157*	*218*	375		**3.546687425**
	133.797054	**241.202946**			**2.23204862**
	218	393	611	critical value=	**16.1963657**
				phi=	**0.162812664**

Table 10 Contingency Table comparing Presenting/Presuming Reference with position in Theme/Rheme – All selected texts

	Presenting	Presuming			χ^2 table
Theme	*123*	*333*	456		**12.55373523**
	169	**286.92981**			**7.58212353**
Rheme	*326*	*429*	755		**7.39714845**
	279.9298101	**475.0701899**			**4.467681713**
	449	762	1211	critical value=	**32.00068892**
				phi=	**0.162557717**

Table 11 Contingency Table comparing Presenting/Presuming Reference with position in Theme/Rheme – All texts

	Presenting	Presuming			χ^2 table
Theme	*228*	*453*	681		**15.68956792**
	296.166937	**384.833063**			**9.132133125**
Rheme	*577*	*593*	1170		**12.07466747**
	508.8330632	**661.1669368**			**7.028075684**
	805	1046	1851	critical value=	**43.92444421**
				phi=	**0.154045823**

Table 12 Contingency Table comparing Nominal/Pronominal Presumed Reference with Theme/Rheme position

	Theme	Rheme			χ^2 table
Presumed-variable Nominal	*438*	*592*	1030		**3.994209**
	481.8713	**548.1287**			**24.63494**
Presumed-variable Pronominal	*122*	*45*	167		**3.511393**
	78.12865	**88.87135**			**21.65709**
	560	637	1197	critical value=	**53.79764**
				phi=	**0.211999**

Table 13 Contingency Table comparing Directed/ Undirected Pronominal Presumed Reference with Theme/Rheme position

	Theme	Rheme			χ^2 table
Presumed-variable Pronominal Undirected	*400*	*579*	979		**0.639187**
	416.3126	**562.6874**			**12.26988**
Presumed-variable Pronominal Directed	*38*	*13*	51		**0.472912**
	21.68738	**29.31262**			**9.078056**
	438	592	1030	critical value=	**22.46004**
				phi=	**0.147668**

Appendix 3.2 Combination of Participant Identification and Tracking, Theme and Information Analyses for "Milling Machine" Text

Text	11.1 Milling-machine elements	The main elements of	a typical knee-and-column horizontal milling machine
Tracking	neutralised	generic --t asserting individuated	specified --t unmarked
Identification	no-referent---addition	single within-group---esphora part meronymy-constitution	no-referent---addition
Thematic Function		Topical	
Theme		Theme	

are shown in	Fig. 11.3.	The elements of	a vertical machine
no identification	specified --t unique—i	specified --t asserting -	specified --t unmarked
no tracking	no-referent---addition	single within-group---esphora part meronymy-constitution	no-referent---addition
	clause-complex-final	simple topical	
	Rheme	Theme	

are	the same	except that
no identification	specified general-comparison semblance-identity asserting -	no identification
no tracking	single preceding---anaphora substitution	no tracking
	clause-final	initial textual
	Rheme	Theme

the spindle head	is mounted at	the top	of the column,
generic --t asserting -	no identification	specified --t – asserting	--t specified - asserting
single non-verbal---exophora	no tracking	single within-group---esphora facet meronymy-constitution	single non-verbal---exophora
non-initial topical		clause-internal	clause-final
		Rheme	

as	shown in	Fig. 11.4.	Column
no identification	no identification	specified --t non-particular	neutralised
no tracking	no tracking	no-referent---addition	no-referent---addition
simple textual		clause-complex-final	
Theme		Rheme	

and base	The column and base	form	the foundation of
neutralised	specified --t asserting -	no identification	--t specified - asserting
no-referent---addition	single non-verbal---exophora	no tracking	single within-group---esphora meronymy-constitution part
	simple topical		clause-internal
	Theme		Rheme

the complete machine.	Both	are made from	cast iron,
--t specified - asserting	--t specified asserting inclusive	no identification	--t generic unmarked
single preceding---anaphora experiential	single preceding---anaphora substitution	no tracking	no-referent---addition
clause-complex-final	simple topical		clause-final
	Theme		Rheme

designed with	thick sections	to ensure	complete rigidity and
no identification	--t generic unmarked	no identification	generic --t unmarked
no tracking	no-referent---addition	no tracking	no-referent---addition
	clause-internal		clause-internal

freedom from	vibration.	The base,	upon which
--t generic unmarked	--t generic unmarked	--t specified - asserting	--t specified non-interlocuters
no-referent---addition	no-referent---addition	single preceding---anaphora complete-repetition	single preceding---anaphora substitution
clause-internal	clause-complex-final	simple topical	initial experiential
		Theme	Theme

the column	is mounted,	is also	the cutting-fluid reservoir
--t specified - asserting	no identification	no identification	--t specified - asserting
single preceding---anaphora complete-repetition	no tracking	no tracking	single non-verbal---exophora
topical non-initial			clause-final
			Rheme

and	contains	the pump	to circulate	the fluid to
no identification	no identification	--t specified - asserting	no identification	--t specified - asserting
no tracking	no tracking	single non-verbal---exophora	no tracking	single preceding---anaphora facet meronymy-constitution
textual simple		clause-internal		clause-internal
Theme		Rheme		

the cutting area.	The column	contains
--t specified - asserting	--t specified – asserting	no identification
single non-verbal---exophora	single preceding---anaphora complete-repetition	no tracking
clause-complex-final	simple topical	
	Theme	

the spindle,	accurately located in	precision bearings.	The spindle
--t specified - asserting	no identification	--t generic unmarked	--t specified - asserting
single non-verbal---exophora	no tracking	no-referent---addition	single preceding---anaphora complete-repetition
clause-final		clause-complex-final	simple topical
Rheme			Theme

is driven through	a gearbox from	a vee-belt drive from	the electric motor
no identification	--t specified unmarked	--t specified unmarked	--t specified - asserting
no tracking	no-referent---addition	no-referent---addition	single context-of-culture---homphora
	clause-internal	clause-internal	clause-internal
	Rheme		

housed at	the base of	the column.	The gearbox
no identification	--t specified - asserting	--t specified - asserting	--t specified - asserting
no tracking	single within-group---esphora part meronymy-constitution	single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition
	clause-internal	clause-complex-final	simple topical
	Rheme		Theme

enables	a range of	spindle speeds	to be selected.
no identification	generic --t unmarked	--t generic unmarked	no identification
no tracking	single within-group---esphora measure meronymy-constitution	no-referent---addition	no tracking
	clause-internal	clause-internal	clause-complex-final
	Rheme		

In the model	shown,	twelve spindle speeds from	32 to
--t specified - asserting	no identification	--t specified particular	--t specified unmarked
single non-verbal---exophora	no tracking	single preceding---anaphora complete-repetition	single preceding---anaphora no-poles---cycle
experiential initial		non-initial topical	
Theme			

1400rev/min	are available.	The front of	the column
--t specified unmarked	no identification	--t specified - asserting	--t specified - asserting
single preceding---anaphora no-poles---cycle	no tracking	single within-group---esphora part meronymy-constitution	single preceding---anaphora complete-repetition
	clause-complex-final	simple topical	
	Rheme	Theme	

carries	the guideways	upon which	the knee
no identification	--t specified - asserting	--t specified non-interlocuters	--t specified - asserting
no tracking	no-referent---addition	single preceding---anaphora substitution	single non-verbal---exophora
	clause-final	experiential initial	topical non-initial
	Rheme	Theme	

is located	and	guided in	a vertical direction.
no identification	no identification	no identification	--t specified unmarked
no tracking	no tracking	no tracking	no-referent---addition
clause-final	textual simple		clause-complex-final
	Theme		Rheme

Figure 11.3	Main machine elements of horizontal milling machine	Figure 11.4	Top of column of vertical milling machine
neutralised	neutralised	neutralised	neutralised
no-referent---addition	no-referent---addition	no-referent---addition	no-referent---addition

Knee	The knee,	mounted on	the column guideways,
neutralised	specified --t asserting -	no identification	--t specified - asserting
single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	no tracking	single preceding---anaphora part meronymy-constitution
	simple topical		clause-final
	Theme		Rheme

provides	the vertical movement of	the table.	Power feed
no identification	--t specified - asserting	--t specified - asserting	--t generic unmarked
no tracking	single within-group---esphora experiential	single non-verbal---exophora	no-referent---addition
	clause-internal	clause-complex-final	simple topical
	Rheme		Theme

is available, through	a gearbox	mounted on	the side, from
no identification	--t specified unmarked	no identification	--t specified - asserting
no tracking	no-referent---addition	no tracking	single preceding---anaphora meronymy-constitution part
	clause-internal		
	Rheme		clause-internal

a separate built-in motor,	providing	a range of	twelve feed rates from
specified difference general-comparison unmarked	no identification	specified --t unmarked	--t specified unmarked
no-referent---addition	no tracking	single within-group---esphora measure meronymy-constitution	no-referent---addition
clause-internal		clause-internal	clause-internal

6 to	250 mm/min.	Drive	is through
--t specified unmarked	--t specified unmarked	generic --t unmarked	no identification
single preceding---anaphora no-poles---cycle	single preceding---anaphora no-poles---cycle	no-referent---addition	no tracking
clause-internal	clause-complex-final	simple topical	
		Theme	

a leadscrew,	whose	bottom end	is fixed to
--t specified unmarked	--t specified non-interlocuters	--t specified unmarked	no identification
no-referent---addition	single preceding---anaphora substitution	single preceding---anaphora meronymy-constitution part	no tracking
clause-final	initial experiential	non-initial topical	
Rheme	Theme		

the machine base.	Provision	is made to raise and lower
--t specified - asserting	--t generic unmarked	no identification
single preceding---anaphora derivation	no-referent---addition	no tracking
clause-complex-final	simple topical	
Rheme	Theme	

the knee	by hand	through
--t specified - asserting	--t generic unmarked	no identification
single preceding---anaphora complete-repetition	no-referent---addition	no tracking
clause-internal	clause-internal	
Rheme		

a leadscrew and nut	operated by	a handwheel at	the front.
--t specified unmarked	no identification	--t specified unmarked	--t specified - asserting
no-referent---addition	no tracking	no-referent---addition	single preceding---anaphora meronymy-constitution part
clause-internal		clause-internal	clause-complex-final

The knee	has	guideways on
--t specified - asserting	no identification	generic --t unmarked
single preceding---anaphora complete-repetition	no tracking	no-referent---addition
simple topical		clause-internal
Theme		Rheme

its top surface	giving	full-width support to	the saddle
--t specified non-interlocuters	no identification	--t generic unmarked	--t specified - asserting
single preceding---anaphora meronymy-constitution part	no tracking	no-referent---addition	single non-verbal---exophora
clause-final		clause-internal	clause-final

and	guiding	it in	a transverse direction.
no identification	no identification	--t specified non-interlocuters	specified --t unmarked
no tracking	no tracking	multiple---ambiguous preceding---anaphora substitution	no-referent---addition
simple textual		clause-internal	clause-complex-final
Theme		Rheme	

A lock	is provided to clamp	the knee in	any vertical position on
--t specified unmarked	no identification	--t specified - asserting	--t specified unrestricted-2
no-referent---addition	no tracking	single preceding---anaphora complete-repetition	single within-group---esphora attitudinal
simple topical		clause-internal	clause-internal
Theme		Rheme	

the column.	Saddle	The saddle,
--t specified - asserting	neutralised	specified --t asserting -
single preceding---anaphora complete-repetition	single non-verbal---exophora	single preceding---anaphora complete-repetition
clause-complex-final		simple topical
		Theme

mounted on	the knee guideways,	provides	the transverse movement of
no identification	--t specified - asserting	no identification	--t specified - asserting
no tracking	single preceding---anaphora derivation	no tracking	single within-group---esphora role-reversal---converse
	clause-final		clause-internal
	Rheme		

the table.	Power feed	is provided through
--t specified - asserting	--t generic unmarked	no identification
single preceding---anaphora complete-repetition	no-referent---addition	no tracking
clause-complex-final	simple topical	
	Theme	

the gearbox on	the knee.	A range of
--t specified - asserting	--t specified - asserting	--t specified unmarked

single within-group---esphora meronymy-constitution part	single preceding---anaphora complete-repetition	single within-group---esphora measure meronymy-constitution
clause-internal	clause-complex-final	simple topical
Rheme		Theme

twelve feeds	is available, from	12 to
--t specified unmarked	no identification	--t specified unmarked
no-referent---addition	no tracking	single preceding---anaphora no-poles---cycle
simple topical		clause-internal
		Rheme

500 mm/min.	Alternative hand movement	is provided through	a leadscrew and nut by
--t specified unmarked	specified difference general-comparison unmarked	no identification	--t specified unmarked
single preceding---anaphora no-poles---cycle	single preceding---anaphora non-gradable---complementarity	no tracking	no-referent---addition
clause-complex-final	simple topical		clause-internal
	Theme		Rheme

a handwheel at	the front of	the knee.
--t specified unmarked	--t specified - asserting	--t specified - asserting
no-referent---addition	single within-group---esphora meronymy-constitution part	single preceding---anaphora complete-repetition
clause-internal	clause-internal	clause-complex-final

Clamping of	the saddle to	the knee
--t generic unmarked	--t specified - asserting	--t specified - asserting
single within-group---esphora enhancement	single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition
simple topical	simple topical	simple topical
Theme		

is achieved by	two clamps on	the side of
no identification	--t specified unmarked	--t specified - asserting
no tracking	single preceding---anaphora inflexion	single preceding---anaphora meronymy-constitution part
	clause-internal	clause-internal
	Rheme	

the saddle.	The saddle	has
--t specified - asserting	--t specified - asserting	no identification
single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	no tracking
clause-complex-final	simple topical	
	Theme	

dovetail guideways on	its upper surface,	at right angles to
--t specified unmarked	--t specified non-interlocuters	no identification
no-referent---addition	single preceding---anaphora meronymy-constitution part	no tracking
clause-internal	clause-internal	clause-internal
Rheme		

the knee guideways,	to	provide	a guide to
--t specified - asserting	no identification	no identification	--t generic unmarked
single preceding---anaphora derivation	no tracking	no tracking	no-referent---addition
clause-final	simple textual		clause-internal
	Theme		Rheme

the table in	a longitudinal direction.	Table
--t specified - asserting	--t specified unmarked	neutralised
single preceding---anaphora complete-repetition	multiple---ambiguous preceding---anaphora experiential	single preceding---anaphora complete-repetition
clause-internal	clause-complex-final	

The table	provides	the surface
specified --t asserting -	no identification	specified --t asserting -
single preceding---anaphora complete-repetition	no tracking	no-referent---addition
simple topical		clause-internal
Theme		Rheme

upon which	all workpieces and workholding equipment	are located and clamped.
--t specified non-interlocuters	specified --t total-nominal	no identification
single preceding---anaphora substitution	no-referent---addition	no tracking
clause-internal	clause-internal	clause-complex-final

A series of	tee slots	is provided for
--t specified unmarked	--t specified unmarked	no identification
single within-group---esphora meronymy-constitution measure	no-referent---addition	no tracking
simple topical	simple topical	
Theme		

this purpose.	The dovetail guides on	the undersurface
--t specified - asserting	--t specified - asserting	--t specified - asserting
single preceding---anaphora attitudinal	single within-group---esphora meronymy-constitution part	single preceding---anaphora meronymy-constitution part
clause-complex-final	simple topical	simple topical
Rheme	Theme	

locate in	the guideways on	the saddle,
no identification	--t specified - asserting	--t specified - asserting
no tracking	single within-group---esphora part meronymy-constitution	single preceding---anaphora complete-repetition
	clause-internal	clause-final

giving	straight-line movement to	the table in
no identification	--t specified unmarked	--t specified - asserting
no tracking	multiple---ambiguous preceding---anaphora experiential	single preceding---anaphora complete-repetition
	clause-internal	clause-internal

a longitudinal direction at right angles to	the saddle movement.	Power feed
--t specified unmarked	--t specified - asserting	--t generic unmarked
no-referent---addition	single preceding---anaphora experiential	no-referent---addition
clause-internal	clause-complex-final	simple topical
		Theme

is provided from	the knee gearbox,	through the saddle,
no identification	--t specified - asserting	--t specified - asserting
no tracking	single preceding---anaphora experiential	single preceding---anaphora complete-repetition
	clause-internal	clause-internal

to the table leadscrew.	Alternative hand feed
--t specified - asserting	--t specified unmarked
single preceding---anaphora meronymy-constitution part	single preceding---anaphora non-gradable---complementarity
clause-complex-final	simple topical
	Theme

is provided by	a handwheel at	each end of
no identification	--t specified unmarked	--t specified asserting individuated
no tracking	no-referent---addition	single within-group---esphora meronymy-constitution part
	clause-internal	clause-internal
	Rheme	

the table.	Stops at	the front of
--t specified - asserting	--t generic unmarked	--t specified - asserting
single preceding---anaphora complete-repetition	no-referent---addition	single preceding---anaphora meronymy-constitution part
clause-complex-final	simple topical	simple topical
	Theme	

the table	can be set to disengage	the longitudinal feed

--t specified - asserting	no identification	--t specified - asserting
single preceding---anaphora complete-repetition	no tracking	single preceding---anaphora experiential
simple topical		clause-internal
		Rheme

automatically in	each direction.	Spindle	The spindle,
no identification	--t specified asserting individuated	neutralised	specified --t asserting -
no tracking	single preceding---anaphora role-reversal---converse	no-referent---addition	single preceding---anaphora complete-repetition
clause-internal	clause-complex-final		simple topical
			Theme

accurately mounted in	precision bearings,	provides	the drive for
no identification	--t specified unmarked	no identification	--t specified - asserting
no tracking	no-referent---addition	no tracking	single within-group---esphora facet meronymy-relational
clause-internal	clause-final		clause-internal
Rheme			Rheme

the milling cutters.	Cutters	can be mounted straight on
--t specified - asserting	--t generic - asserting	no identification
no-referent---addition	single preceding---anaphora complete-repetition	no tracking
clause-complex-final	simple topical	
	Theme	

the spindle nose	or in cutterholding devices	which
--t specified - asserting	--t generic - asserting	--t specified non-interlocuters
single preceding---anaphora meronymy-constitution part	no-referent---addition	single preceding---anaphora substitution
clause-internal		initial topical
Rheme	clause-final	Theme

in turn	are mounted in	the spindle,	held in
no identification	no identification	--t specified - asserting	no identification
no tracking	no tracking	single preceding---anaphora complete-repetition	no tracking
non-initial textual		clause-final	
		Rheme	

position by	a drawbolt	passing through	the hollow spindle.
generalized	--t specified unmarked	no identification	--t specified - asserting
single preceding---anaphora elaboration	no-referent---addition	no tracking	single preceding---anaphora substitution
clause-internal	clause-internal	clause-internal	clause-complex-final

Spindles of	milling machines	have
--t generic – asserting	--t generic - asserting	no identification
single preceding---anaphora inflexion	single preceding---anaphora inflexion	no tracking
simple topical	simple topical	

Theme		

a standard spindle nose,	as	shown in	Fig. 11.5,
--t specified unmarked	no identification	no identification	--t specified unmarked
single preceding---anaphora substitution	no tracking	no tracking	single non-verbal---exophora
clause-final	simple textual		clause-final
Rheme	Theme		Rheme

to	allow for	easy interchange of	cutters and
no identification	no identification	--t specified unmarked	--t generic - asserting
no tracking	no tracking	no-referent---addition	single preceding---anaphora complete-repetition
simple textual		clause-internal	clause-internal
Theme		Rheme	

cutter-holding devices.	The bore of	the nose
--t generic - asserting	--t asserting - generic	--t specified - asserting
single preceding---anaphora complete-repetition	single within-group---esphora facet meronymy-constitution	single preceding---anaphora complete-repetition
clause-complex-final	simple topical	simple topical
	Theme	

is tapered to provide	accurate location,	the angle of
no identification	--t generic unmarked	--t specified - asserting
no tracking	no-referent---addition	single within-group---esphora meronymy-constitution measure
	clause-final	simple topical
	Rheme	Theme

taper	being	16° 36'.
--t generic unmarked	no identification	--t specified unmarked
no-referent---addition	no tracking	single preceding---anaphora no-poles---cycle
simple topical		clause-complex-final
Theme		Rheme

The diameter of	the taper	depends on
--t specified - asserting	--t specified - asserting	no identification
single within-group---esphora meronymy-constitution measure	single preceding---anaphora complete-repetition	no tracking
simple topical	simple topical	
Theme		

the size of	the machine	and
--t specified - asserting	--t specified - asserting	no identification
single within-group---esphora meronymy-constitution measure	single preceding---anaphora substitution	no tracking
clause-internal	clause-final	simple textual
Rheme		Theme

may be	30, 40, or 50 IST (International Standard Taper).	Due to	their steepness of
no identification	--t specified unmarked	no identification	--t specified non-interlocuters
no tracking	single preceding---anaphora no-poles---cycle	no tracking	single beyond-group---cataphora meronymy-constitution facet
	clause-complex-final	initial experiential	initial experiential
		Theme	

angle,	these tapers	- known as
--t generic - asserting	--t specified proximate	no identification
single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	no tracking
initial experiential	non-initial topical	

non-stick or	self-releasing	- cannot be relied upon to transmit
--t generic unmarked	--t generic unmarked	no identification
no-referent---addition	no-referent---addition	no tracking
clause-internal	clause-final	
Rheme		

the drive to	the cutter or	cutter-holding device.
--t specified - asserting	--t specified - asserting	--t specified - asserting
single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition
clause-internal	clause-internal	clause-complex-final
Rheme		

Two driving keys	are provided to transmit	the drive.
--t specified unmarked	no identification	--t specified - asserting
no-referent---addition	no tracking	single preceding---anaphora complete-repetition
simple topical		clause-complex-final
Theme		Rheme

Figure 11.5	Standard milling-machine spindle nose	Cutters
neutralised	neutralised	specified --t unmarked
no-referent---addition	no-referent---addition	single preceding---anaphora complete-repetition
		simple topical
		Theme

which	are mounted directly on	the spindle nose
--t specified non-interlocuters	no identification	--t specified - asserting
single preceding---anaphora substitution	no tracking	single preceding---anaphora complete-repetition
simple topical		clause-final
Theme		Rheme

are located on	a centring arbor,	and	four tapped holes
no identification	--t specified unmarked	no identification	--t specified unmarked
no tracking	no-referent---addition	no tracking	no-referent---addition
	clause-final	initial textual	non-initial topical
	Rheme	Theme	Rheme

are provided to hold	the cutter	in position.
no identification	--t specified – asserting	generalized
no tracking	single preceding---anaphora complete-repetition	single preceding---anaphora elaboration
	clause-internal	clause-complex-final
	Rheme	

The two keys	again	provide	the means of
--t specified - asserting	no identification	no identification	generic --t - asserting
single preceding---anaphora substitution	no tracking	no tracking	single within-group---esphora elaboration
initial topical	non-initial textual		clause-internal
Theme			Rheme

transmitting	the drive.	The spindle of
--t generic - asserting	--t specified - asserting	--t specified - asserting
single preceding---anaphora derivation	single preceding---anaphora complete-repetition	single within-group---esphora meronymy-constitution part
clause-internal	clause-complex-final	simple topical
		Theme

a horizontal machine	is	fixed	and	cannot be adjusted in
--t specified unmarked	no identification	no identification	no identification	no identification
no-referent---addition	no tracking	no tracking	no tracking	no tracking
simple topical		clause-final	simple textual	
		Rheme	Theme	

an axial direction,	i.e.	along its axis.
--t specified unmarked	no identification	--t specified non-interlocuters
single preceding---anaphora experiential	no tracking	no-referent---addition
clause-internal	clause-internal	clause-complex-final
Rheme		

On vertical machines,	provision	is made for
--t specified - asserting	--t generic unmarked	no identification
single preceding---anaphora role-reversal---converse	no-referent---addition	no tracking
initial experiential	non-initial topical	
Theme		

axial movement,	which	is controlled by
--t generic – asserting	--t specified non-interlocuters	no identification
single preceding---anaphora complete-repetition	single preceding---anaphora substitution	no tracking
clause-final	simple topical	
Rheme		

a handwheel on	the spindle head.	The spindle	runs in
--t specified unmarked	--t specified - asserting	--t specified - asserting	no identification
no-referent---addition	single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	no tracking
clause-internal	clause-complex-final	simple topical	
		Theme	

a quill	which	is moved through	a rack and pinion
--t specified unmarked	--t specified non-interlocuters	no identification	--t specified unmarked
no-referent---addition	Single preceding---anaphora substitution	no tracking	no-referent---addition
clause-final	simple topical		clause-internal
Rheme	Theme		Rheme

in the same way as	a drilling machine spindle	(see
specified general-comparison semblance-identity - asserting	--t specified unmarked	no identification
single preceding---anaphora substitution	no-referent---addition	no tracking
clause-internal	clause-internal	simple experiential
		Theme

Fig. 7.2).	A locking bolt	is provided to lock	the quill in
--t specified unmarked	--t specified unmarked	no identification	--t specified - asserting
no-referent---addition	no-referent---addition	no tracking	single preceding---anaphora complete-repetition
clause-complex-final	simple topical		clause-internal
Rheme	Theme		Rheme

any position along	its operating length.	Overarm and arbor support
--t specified unrestricted-2	--t specified non-interlocuters	neutralised
no-referent---addition	single preceding---anaphora meronymy-constitution facet	single non-verbal---exophora
clause-internal	clause-complex-final	

The majority of	cutters	used on
generic --t asserting -	--t generic - asserting	no identification

single within-group---esphora hyponymy---class-preceding superordination-class-subclass	single preceding---anaphora complete-repetition	no tracking
simple topical	simple topical	
Theme		

horizontal machines	are held on	an arbor
--t generic - asserting	no identification	--t specified unmarked
single preceding---anaphora complete-repetition	no tracking	no-referent---addition
simple topical		clause-final
Theme		Rheme

which	is located and held in	the spindle.	Due to
--t specified non-interlocuters	no identification	--t specified - asserting	no identification
single preceding---anaphora substitution	no tracking	single preceding---anaphora complete-repetition	no tracking
simple topical		clause-complex-final	initial experiential
Theme		Rheme	Theme

the length of	the arbors	used,
--t specified - asserting	--t specified - asserting	no identification
single within-group---esphora meronymy-constitution measure	single preceding---anaphora complete-repetition	no tracking
initial experiential	initial experiential	non-initial topical

support	is required at	the outer end	to prevent
--t generic - asserting	no identification	specified difference -- comparison-quality - asserting	no identification
no-referent---addition	no tracking	single preceding---anaphora meronymy-constitution part	no tracking
non-initial topical		clause-internal	
		Rheme	

deflection	when	cutting	takes	place.
--t generic unmarked	no identification	specified --t total-nominal	no identification	specified --t total-nominal
no-referent---addition	no tracking	no-referent---addition	no tracking	single preceding---anaphora elaboration
clause-final	textual initial	topical non-initial		clause-complex-final

Support	is provided by	an arbor-support bracket,
--t generic - asserting	no identification	--t specified unmarked
single preceding---anaphora complete-repetition	no tracking	no-referent---addition
simple topical		clause-final
Theme		Rheme

clamped to	an overarm	which	is mounted on
no identification	--t specified unmarked	--t specified non-interlocuters	no identification
no tracking	no-referent---addition	single preceding---anaphora	no tracking

		substitution	
	clause-final	simple topical	
	Rheme	Theme	

top of	the column in	a dovetail slide.
--t specified - asserting	--t specified - asserting	--t specified unmarked
single within-group---esphora meronymy-constitution part	single preceding---anaphora complete-repetition	no-referent---addition
clause-internal	clause-internal	clause-complex-final
Rheme		

The overarm	is adjustable in or out for	different lengths of
--t specified - asserting	no identification	specified difference -- comparison-quantity unmarked
single preceding---anaphora complete-repetition	no tracking	single within-group---esphora meronymy-constitution measure
simple topical		clause-internal
Theme		Rheme

arbor,	or	can be fully pushed in	when
--t generic - asserting	no identification	no identification	no identification
single preceding---anaphora complete-repetition	no tracking	no tracking	no tracking
clause-final	simple textual	clause-final	initial textual
	Theme	Rheme	Theme

arbor support	is	not required.	Two clamping bolts
--t generic - asserting	no identification	no identification	--t specified unmarked
single preceding---anaphora derivation	no tracking	no tracking	no-referent---addition
non-initial topical		clause-complex-final	simple topical
		Rheme	Theme

are provided to lock	the overarm in	any position.
no identification	--t specified - asserting	--t specified unrestricted-2
no tracking	single preceding---anaphora complete-repetition	no-referent---addition
	clause-internal	clause-complex-final
	Rheme	

The arbor support	is located in	the overarm dovetail	and
--t specified - asserting	no identification	--t specified - asserting	no identification
single preceding---anaphora derivation	no tracking	single preceding---anaphora derivation	no tracking
simple topical		clause-final	simple textual
Theme		Rheme	Theme

is locked	by means of	its clamping bolt.
no identification	no identification	--t specified non-interlocuters
no tracking	no tracking	single preceding---anaphora meronymy-constitution part
	clause-internal	clause-complex-final
	Rheme	

A solid bearing	is provided	in which	the arbor
--t specified unmarked	no identification		--t specified - asserting
no-referent---addition	no tracking		single preceding---anaphora complete-repetition
simple topical		initial experiential	topical non-initial
Theme			

runs	during	spindle rotation
no identification	no identification	generic --t unmarked
no tracking	no tracking	no-referent---addition
	clause-internal	clause-complex-final
	Rheme	

Appendix 3.3 Combination of Participant Identification and Tracking, Theme and Information Analyses for "Broadcast Networks" Text

Text	14.2	BROADCAST NETWORKS	Chapter 12	(section 12.3.2)
Part. ID	no-referent---addition	Neutralised	--t specified unmarked	--t specified unmarked
Part. Track		no-referent---addition	single non-verbal---exophora	single non-verbal---exophora
Th-Rh Type			initial topical	experiential non-initial
Th-Rh			Theme	

introduced	the concept	of broadcast networks	and
	--t specified - asserting	--t specified unmarked	
	single within-group---esphora relational hyponymy---class-preceding	no-referent---addition	
	clause-internal	clause-internal	clause-internal
	Rheme		

the most common topologies.	Remember	that	broadcast networks
specified --t asserting quality			--t specified - asserting
single preceding---anaphora complete-repetition			no-referent---addition
clause-complex-final	simple topical	clause-final	simple topical
	Theme	Rheme	Theme

use	a channel	to which	all the users
	--t specified unmarked	--t specified non-interlocuters	--t specified - asserting
	no-referent---addition	single preceding---anaphora substitution	single preceding---anaphora part meronymy-constitution
	clause-final	textual initial	topical non-initial
	Rheme	Theme	

are	connected,	so	all the users	receive
			--t specified - asserting	
			single preceding---anaphora complete-repetition	
	clause final	textual initial	topical non-initial	
		Theme		

any transmission	made	on the channel.	The only wide area broadcast networks
specified --t unrestricted-2		--t specified - asserting	--t specified - asserting
no-referent---addition		single preceding---anaphora complete-repetition	single preceding---anaphora derivation
clause-internal		clause-complex-final	topical initial
Rheme			Theme

all	use	radio broadcast,	either relatively local
		--t specified unmarked	
		no-referent---addition	
non-initial experiential		clause-internal	clause-internal
		Rheme	

up to a few hundred kilometres	or over much further distances	using	satellite transmission.
--t specified unmarked	specified difference -- comparison-quality unmarked		--t specified unmarked
no-referent---addition	single preceding---anaphora experiential		no-referent---addition
clause-internal	clause-internal	clause-complex-final	

Section 14.2.1	will examine	this	further.
--t specified unmarked		specified --t non-interlocuters	
single non-verbal---exophora		single preceding---anaphora substitution	
simple topical		clause-internal	clause-complex-final
Theme		Rheme	

The most common examples of	broadcast networks	are	local area networks.
specified --t asserting quality	--t specified unmarked		--t specified unmarked
single within-group---esphora relational hyponymy---class-preceding	no-referent---addition		no-referent---addition
simple topical			clause-complex-final
Theme			Rheme

These	may use	twisted pair cables,	coaxial cable
--t specified non-interlocuters		--t specified unmarked	--t specified unmarked
single preceding---anaphora substitution		no-referent---addition	no-referent---addition
simple topical		clause-internal	clause-internal
Theme		Rheme	

or optical fibre cable	as	their transmission medium	(see
--t specified unmarked		--t specified - asserting	
no-referent---addition		single preceding---anaphora meronymy-constitution part	
clause-internal	clause-final		simple topical
			Theme

section 13.6).	Comparison	between twisted pair	and coaxial cable
specified --t unmarked	specified --t unmarked	generic --t unmarked	--t specified unmarked
single non-verbal---exophora	no-referent---addition	single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition
clause-complex-final	simple topical		
Rheme	Theme		

is not	helpful	because	there	are
	clause-final	textual initial	topical non-initial	
	Rheme	Theme		

many variants	of each	to meet	the different requirements
specified --t non-particular	specified --t asserting individuated		--t specified - asserting
single within-group---esphora relational hyponymy---class-preceding	single preceding---anaphora substitution		single within-group---esphora enhancement
clause-internal	clause-final		clause-internal
Rheme			Rheme

of bandwidth,	loss,	noise immunity,	etc.
--t specified unmarked	--t specified unmarked	--t specified unmarked	
no-referent---addition	no-referent---addition	no-referent---addition	multiple---ambiguous context-of-culture---homophora
clause-internal	clause-internal	clause-internal	clause-complex-final

In general,	coaxial cable	has	higher noise immunity
	--t specified unmarked		specified unmarked difference general-comparison
	no-referent---addition		no-referent---addition
initial textual	topical non-initial		clause-internal
Theme			Rheme

and bandwidth	but	the cable	is stiffer
specified difference -- comparison-quantity unmarked		specified --t asserting -	
no-referent---addition		single preceding---anaphora derivation	
clause-final	textual initial	topical non-initial	clause-final
	Theme		Rheme

(which	may or may not be	helpful	depending on
textual simple		clause-final	
Theme		Rheme	

whether	it	is being	surface mounted	or
	--t specified non-interlocuters			
	single preceding---anaphora substitution			
initial textual	non-initial topical		clause final	simple textual
Theme			Rheme	Theme

pushed through	ducts).	Both types	can adequately serve
	--t specified unmarked	specified --t asserting inclusive	
	no-referent---addition	single preceding---anaphora relational hyperonymy---subclass-preceding	
	clause-complex-final	simple topical	
	Rheme	Theme	

most LAN environments,	but	coaxial cable	has been dropping out of
--t specified asserting quality		--t specified - asserting	
no-referent---addition		single preceding---anaphora complete-repetition	
clause-final	textual initial	topical non-initial	
Rheme	Theme		

use.	Optical fibre	is particularly suited to	environments
--t generic - asserting	--t specified unmarked		--t specified unmarked
no-referent---addition	single preceding---anaphora complete-repetition		single preceding---anaphora derivation
clause-complex-final	simple topical		clause-final
Rheme	Theme		Rheme

which	have	high levels	of electromagnetic

			radiation,
--t specified non-interlocuters		--t specified unmarked	--t specified unmarked
single preceding---anaphora substitution		single within-group---esphora facet meronymy-constitution	no-referent---addition
simple textual		clause-internal	clause-final
Theme		Rheme	

or	to meet	demands	for very high speeds
		--t specified unmarked	--t specified unmarked
		single within-group---esphora elaboration	single within-group---esphora meronymy-constitution facet
textual simple		clause-internal	clause-internal
Theme		Rheme	

of transmission.	However,	it	is more difficult to	tap into,
--t specified unmarked		--t specified non-interlocuters		
no-referent---addition		single preceding---anaphora substitution		
clause-complex-final	textual initial	topical non-initial		clause-final
	Theme			Rheme

which	makes	it	more difficult and expensive
		specified --t non-interlocuters	
		single preceding---anaphora substitution	
textual simple		clause-internal	clause-internal
Theme		Rheme	

for the installation	of a LAN.	LAN protocols	have developed into
--t specified – asserting	--t specified unmarked	--t specified unmarked	
single within-group---esphora elaboration	no-referent---addition	no-referent---addition	
clause-internal	clause-complex-final	simple topical	
		Theme	

the following layers:	* physical layer	identical to	ISO layer 1
--t specified – asserting	--t specified unmarked		--t specified unmarked
single beyond-group---cataphora hyponymy---class-preceding relational	single preceding---anaphora hyponymy---class-preceding superordination-class-subclass		no-referent---addition
clause-final	clause-internal		clause-final
Rheme			Rheme

* medium access control	(MAC)	layer	to manage
--t specified	--t specified unmarked	--t specified unmarked	

unmarked			
no-referent---addition	single preceding---anaphora substitution	single preceding---anaphora superordination-class-subclass hyponymy---class-preceding	
clause-internal	clause-internal	clause-internal	

communications	over the link	* logical link control	(LLC)
--t unmarked generic	--t specified - asserting	--t specified unmarked	--t specified unmarked
no-referent---addition	multiple---ambiguous context-of-culture---homophora	no-referent---addition	single preceding---anaphora substitution
clause-internal	clause-final	clause-internal	clause-internal
Rheme			

layer,	which	provides	a form
--t specified unmarked	--t specified non-interlocuters		--t specified unmarked
single preceding---anaphora superordination-class-subclass hyponymy---class-preceding	single preceding---anaphora substitution		single within-group---esphora hyponymy---class-preceding relational
clause-final	textual simple		clause-internal
	Theme		Rheme

of multiplexing	to handle	multiple-source data	(a number
--t specified unmarked		--t generic unmarked	--t generic unmarked
no-referent---addition		no-referent---addition	single within-group---esphora hyponymy---class-preceding relational
clause-internal		clause-internal	clause-internal
		Rheme	

of users	attached to	one host).	In addition,
--t generic unmarked		--t specified particular	
no-referent---addition		no-referent---addition	
clause-internal		clause-complex-final	textual initial
		Rheme	Theme

the LLC layer	assembles	the data	into a frame
--t specified - asserting		--t specified – asserting	--t specified unmarked
single preceding---anaphora complete-repetition		single preceding---anaphora complete-repetition	no-referent---addition
topical non-initial		clause-internal	clause-internal
		Rheme	

complete with	address	and error checking bits	and
	--t generic unmarked	--t generic unmarked	
	no-referent---addition	no-referent---addition	
	clause-internal	clause-final	textual simple
	Rheme		Theme

disassembles	them	on receipt.	For a particular LLC protocol
	--t specified non-interlocuters	--t generic - asserting	--t specified particular
	single preceding---anaphora substitution	no-referent---addition	no-referent---addition
	clause-internal	clause-complex-final	experiential initial
	Rheme		Theme

there	may be	several different MAC options	provided,
		specified difference general-comparison particular	
		single preceding---anaphora complete-repetition	
topical non-initial		clause-final	
		Rheme	

since	this	is	the protocol layer
	--t specified non-interlocuters		--t specified – asserting
	single preceding---anaphora substitution		single preceding---anaphora complete-repetition
textual initial	topical non-initial		clause-final
Theme			Rheme

in which	the differences	in topology	are	involved
--t specified non-interlocuters	specified difference general-comparison - asserting	--t generic non-particular		
single preceding---anaphora substitution	single within-group---esphora hyponymy---class-preceding relational	no-referent---addition		
textual initial	topical non-initial	topical non-initial		clause-complex-final
Theme				Rheme

The major standards activity	for LAN networks	has been developed by	the US Institute of Electrical and Electronic Engineers
--t specified - asserting	--t generic non-particular		--t specified unique--i
single within-group---esphora complete-repetition	single preceding---anaphora complete-repetition		single context-of-culture---homophora
simple topical			clause-internal
Theme			Rheme

(IEEE).	Their work	has been organised into	a number
--t specified unique--i	--t specified - asserting		--t generic unmarked
single context-of-culture---homophora	single preceding---anaphora extension		single within-group---esphora hyponymy---class-preceding relational
clause-complex-final	simple topical		clause-internal

	Theme		Rheme

of committees,	of which	some	are
--t generic non-particular	--t specified non-interlocuters	specified --t restricted	
no-referent---addition	single preceding---anaphora substitution	multiple---ambiguous beyond-group---cataphora substitution	
clause-final	textual initial	topical non-initial	
	Theme		

as follows:	* 802.2	Logical link control (LLC)	* 802.3
	specified --t unique--i	--t specified non-particular	--t specified unique--i
	no-referent---addition	single preceding---anaphora complete-repetition	no-referent---addition
clause-final	clause-internal	clause-internal	clause-internal
Rheme			

CSMA/CD networks	(Ethernet, etc.)	* 802.5	Token ring networks
--t specified non-particular	--t specified unique--i	--t specified unique--i	specified --t unmarked
no-referent---addition	single context-of-culture---homophora	no-referent---addition	no-referent---addition
clause-internal	clause-internal	clause-internal	clause-complex-final

Appendix 3.4 Combination of Participant Identification and Tracking, Theme and Information Analyses for "Security" Text

7.1	Introduction	Security measures	must be incorporated into
	neutralised	specified --t - asserting	
	no-referent---addition	no-referent---addition	
		simple topical	
		Theme	

computer systems	whenever	they
specified --t - asserting		specified --t non-interlocuters
no-referent---addition		multiple---ambiguous preceding---anaphora substitution
clause-final	initial textual	non-initial topical
Rheme	Theme	

are	potential targets	for malicious or mischievous attacks.
	specified --t non-particular	specified --t non-particular
	no-referent---addition	no-referent---addition
	clause-internal	clause-complex-final
	Rheme	

This	is	especially so	for systems
--t specified proximate			specified --t - asserting
multiple---ambiguous preceding---anaphora substitution			no-referent---addition
simple topical		clause-internal	clause-final
Theme		Rheme	

that	handle	financial transactions	or confidential, classified or other information
		specified --t non-particular	specified --t non-particular
		no-referent---addition	no-referent---addition
simple topical		clause-internal	clause-internal
Theme		Rheme	

whose secrecy	and integrity	are	critical.
--t specified - asserting	specified --t asserting -		
single preceding---anaphora alienable meronymy-constitution	single preceding---anaphora alienable meronymy-constitution		
clause-internal	clause-internal		clause-complex-final

In Figure 7.1,	we	summarize
specified --t non-particular	specified --t interlocuters	
single non-verbal---exophora	single non-verbal---exophora	
initial experiential	non-initial topical	
Theme		

the evolution	of security needs	in computer systems
--t specified - asserting	specified --t non-particular	specified --t - asserting
single within-group---esphora facet meronymy-constitution	no-referent---addition	no-referent---addition
clause-internal	clause-internal	clause-final
Rheme		

since	they	first
	specified --t non-interlocuters	
	single preceding---anaphora substitution	
textual initial	topical non-initial	non-initial interpersonal
Theme		

arose	with the advent	of shared data
	--t specified - asserting	specified --t - asserting
	single within-group---esphora alienable meronymy-constitution	no-referent---addition
	clause-internal	clause-internal
	Rheme	

in multi-user timesharing systems	of the 1960s and 70s.	Today
specified --t non-particular	--t specified - asserting	specified --t asserting -
no-referent---addition	context-of-culture---homophora single	context-of-culture---homophora single
clause-internal	clause-complex-final	textual initial
		Theme

the advent	of wide-area, open distributed systems	has resulted in
--t specified - asserting	specified --t non-particular	
single within-group---esphora alienable meronymy-constitution	no-referent---addition	
topical non-initial	topical non-initial	

a wide range	of security issues.	The need
--t specified unmarked	specified --t - asserting	specified --t asserting -
single within-group---esphora relational hyponymy---class-preceding	no-referent--- addition	single within-group--- esphora elaboration
clause-internal	clause-complex-final	simple topical
Rheme		Theme

to protect	the integrity	and privacy
	specified --t asserting -	--t specified - asserting
	single within-group---esphora alienable meronymy-constitution	single within-group---esphora alienable meronymy-constitution
simple topical	simple topical	simple topical

of information	and other resources	belonging to
--t specified - asserting	specified difference general-comparison - asserting	
no-referent---addition	multiple---ambiguous preceding---anaphora attitudinal	
simple topical	simple topical	simple topical

individuals	and organizations	is pervasive	in both the physical
specified --t - asserting	specified --t - asserting		specified --t asserting inclusive
no-referent--- addition	no-referent--- addition		context-of-culture---homophora single
simple topical	simple topical		clause-internal
			Rheme

and the digital world.	It	arises from
--t specified - asserting	--t specified non-interlocuters	
context-of-culture---homophora single	single preceding---anaphora substitution	
clause-complex-final	simple topical	
	Theme	

the desire	to share	resources.
specified --t asserting -		specified --t asserting -
single within-group---esphora elaboration		single context-of-culture--- homophora
clause-internal	clause-internal	clause-complex-final
Rheme		

In the physical world,	organizations	adopt
specified --t asserting -	--t specified - asserting	
single preceding---anaphora no-poles---cycle	single preceding---anaphora complete-repetition	
experiential initial	topical non-initial	
Theme		

security policies	that	provide for
--t specified - asserting	--t specified non-interlocuters	
single preceding---anaphora complete-repetition	single preceding---anaphora substitution	
clause-final	simple topical	
Rheme	Theme	

the sharing	of resources	within specified limits.
--t specified - asserting	--t specified - asserting	--t specified non-particular
single preceding---anaphora derivation	single preceding---anaphora complete-repetition	single preceding---anaphora enhancement
clause-internal	clause-internal	clause-complex-final
Rheme		

For example,	a company	may permit	entry to
generalized	--t specified unmarked		--t specified unmarked
	no-referent---addition		no-referent---addition
textual initial	topical non-initial		clause-internal
Theme			Rheme

its buildings	for its employees	and for accredited visitors.
specified --t - asserting	--t specified - asserting	--t specified unmarked
single preceding---anaphora meronymy-constitution part	single preceding---anaphora part meronymy-constitution	single preceding---anaphora meronymy-constitution part
clause-internal	clause-internal	clause-complex-final

A security policy	for documents	may specify	groups of
--t specified unmarked	--t specified unmarked		specified --t - asserting
no-referent---addition	no-referent---addition		single within-group---esphora relational hyponymy---class-preceding
simple topical	simple topical		clause-internal
Theme			Rheme

employees	who	can access
--t specified unmarked		
single preceding---anaphora complete-repetition		
clause-final	simple topical	
	Theme	

classes	of documents	or
generic --t asserting -	specified --t	

	asserting -	
single within-group---esphora relational hyponymy---class-preceding	no-referent---addition	
clause-internal	clause-final	textual initial
Rheme		Theme

it	may be defined for	individual documents
specified --t non-interlocuters		generic --t asserting -
single preceding---anaphora substitution		single preceding---anaphora derivation
topical non-initial		clause-internal
		Rheme

and users.	Security policies	are enforced with
generic --t asserting -	--t generic - asserting	
single preceding---anaphora experiential	single preceding---anaphora complete-repetition	
clause-complex-final	simple topical	
	Theme	

the help	of security mechanisms.	For example,
--t specified - asserting	--t specified unmarked	--t generic - asserting
single within-group---esphora elaboration	no-referent---addition	no-referent---addition
clause-internal	clause-complex-final	textual initial
Rheme		Theme

access to	a building	may be controlled by
--t specified - asserting	--t specified unmarked	
single within-group---esphora extension	no-referent---addition	
topical non-initial	topical non-initial	

a reception clerk,	who	issues
--t specified unmarked	--t specified non-interlocuters	
no-referent---addition	single preceding---anaphora substitution	
clause-final	simple topical	
Rheme	Theme	

badges	to accredited visitors,	and	enforced by
specified --t unmarked	--t specified unmarked		
no-referent---addition	no-referent---addition		
clause-internal	clause-final	textual simple	
Rheme		Theme	

a security guard	or by electronic door locks.	Access to
--t specified unmarked	specified --t unmarked	--t generic - asserting
no-referent---addition	no-referent---addition	single within-group---esphora extension
clause-internal	clause-complex-final	simple topical
Rheme		Theme

paper documents	is usually controlled	by concealment
--t specified unmarked		--t generic - asserting
no-referent---addition		context-of-culture---homophora single
simple topical		clause-internal
		Rheme

and restricted distribution.	In the electronic world,	the distinction
--t generic - asserting	--t specified - asserting	--t specified - asserting
context-of-culture---homophora single	single preceding---anaphora experiential	single within-group---esphora elaboration
clause-complex-final	initial experiential	topical non-initial
	Theme	

between security policies	and mechanisms	remains	important;
--t specified unmarked	--t specified unmarked		
no-referent---addition	no-referent---addition		
topical non-initial	topical non-initial		clause-final
			Rheme

without it,	it	would be	difficult	to determine
--t specified non-interlocuters	generalized			
single preceding---anaphora substitution				
initial experiential	simple topical		clause internal	clause final
Theme			Rheme	

whether	a particular system	was	secure.	Security policies
	--t specified particular			--t generic - asserting
	no-referent---addition			single preceding---anaphora complete-repetition
textual initial	topical non-initial		clause-complex-final	simple topical
Theme			Rheme	Theme

are	independent of	the technology	used,
		--t specified - asserting	
		context-of-culture---homophora single	
	clause-internal	clause-final	
	Rheme		

just as	the provision	of a lock
	--t specified - asserting	--t specified unmarked
	single within-group---esphora enhancement	no-referent---addition
textual initial	topical non-initial	topical non-initial
Theme		

on a door	does not ensure	the security
--t specified unmarked		--t specified - asserting
no-referent---addition		single preceding---anaphora alienable meronymy-constitution
topical non-initial		clause-internal
		Rheme

of a building	unless	there	is	a policy
--t specified unmarked		generalized		--t specified unmarked
no-referent---addition				no-referent---addition
clause-final	textual initial	topical non-initial		clause-internal
	Theme			Rheme

for its use	(for example,
--t specified - asserting	--t specified unmarked
single preceding---anaphora superordination-class-subclass hyperonymy---subclass-preceding	no-referent---addition
clause-final	clause-internal

that	the door	will be locked
	--t generic - asserting	
	single preceding---anaphora complete-repetition	
clause-internal		

whenever	nobody	is guarding	the entrance).
	generalized		--t generic - asserting
			single preceding---anaphora meronymy-constitution part
textual initial	topical non-initial		clause-complex-final
Theme			Rheme

The security mechanisms	that	we
--t specified - asserting		--t specified interlocuters
single preceding---anaphora complete-repetition		context-of-culture---homophora single
simple topical		simple topical

shall describe	do not	in themselves	ensure
		--t specified non-interlocuters	
		single preceding---anaphora substitution	
		clause-internal	
		Rheme	

the security	of a system.	In Section 7.1.2,
--t specified - asserting	--t generic - asserting	specified --t non-particular
single within-group---esphora alienable meronymy-constitution	no-referent---addition	single non-verbal---exophora
clause-internal	clause-complex-final	experiential initial
		Theme

we	outline	the requirements
--t specified interlocuters		--t specified - asserting
context-of-culture---homophora single		single elaboration within-group---esphora
topical non-initial		clause-internal
		Rheme

for security	in various simple electronic commerce scenarios,	illustrating
--t generic - asserting	specified --t non-particular	
single preceding---anaphora complete-repetition	no-referent---addition	
clause-internal	clause-final	

the need	for policies	in that context.
--t specified - asserting	specified --t unmarked	--t specified distant
single elaboration within-group---esphora	single within-group---esphora alienable meronymy-constitution	single preceding---anaphora experiential
clause-internal	clause-internal	clause-complex-final

As an initial example,	consider	the security
--t specified unmarked		--t specified - asserting
no-referent---addition		single preceding---anaphora complete-repetition
initial textual	non-initial topical	clause-internal
Theme		Rheme

of a networked file server	whose interface	is accessible
--t specified unmarked	--t specified non-interlocuters	
no-referent---addition	single preceding---anaphora alienable meronymy-constitution	
clause-final	simple topical	
	Theme	

to clients.	To ensure that	access control
--t specified unmarked		--t specified - asserting
no-referent---addition		single within-group---esphora extension
clause-complex-final	simple topical	simple topical
Rheme	Theme	

to files	is maintained,	there
--t specified - asserting		
single preceding---anaphora complete-repetition		
simple topical		simple topical
		Theme

would need to be	a policy	that	all requests
	--t specified unmarked		--t specified total-nominal
	no-referent---addition		no-referent---addition
	clause-final	textual initial	topical non-initial
	Rheme	Theme	

must include	an authenticated user identity.	The provision
	--t specified unmarked	--t specified - asserting
	no-referent---addition	single within-group---esphora extension
	clause-complex-final	topical initial
	Rheme	Theme

of mechanisms	for the protection	of data
--t generic - asserting	--t specified - asserting	--t generic - asserting
single within-group---esphora superordination-class-subclass hyponymy---class-preceding	single within-group---esphora extension	no-referent---addition
topical initial	topical initial	topical initial

and other computer- based resources	and for	securing
generic general-comparison difference - asserting		--t specified - asserting
single preceding---anaphora attitudinal		single within-group---esphora extension
topical initial		topical non-initial

networked transactions	is	the concern
--t specified unmarked		--t specified - asserting
no-referent---addition		single within-group---esphora alienable meronymy-constitution
topical non-initial		clause-internal
		Rheme

of this chapter.	We	shall describe
--t specified proximate	--t specified interlocuters	
single non-verbal---exophora	single non-verbal---exophora	
clause-complex-final	simple topical	
	Theme	

the mechanisms	that	enable
--t specified - asserting		
multiple---ambiguous within-group---esphora extension		
clause-final	simple topical	
Rheme	Theme	

security policies	to be enforced	in distributed systems.
specified --t unmarked		--t specified unmarked
no-referent---addition		no-referent---addition
clause-internal		clause-complex-final
Rheme		Rheme

The mechanisms	we	shall describe
--t specified - asserting	--t specified interlocuters	
single preceding---anaphora complete-repetition	context-of-culture---homophora single	
simple topical	simple topical	
Theme		

are	strong enough	to resist	the most determined attacks.
			--t specified asserting quality
			single preceding---anaphora no-poles---cycle
	clause-internal		clause-complex-final
	Rheme		

The distinction between	security policies	and security mechanisms
--t specified - asserting	--t generic - asserting	--t generic - asserting
single within-group---esphora elaboration	single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition
simple topical	simple topical	simple topical
Theme		

is	helpful	when	designing	secure systems,
				--t specified unmarked
				no-referent---addition
	clause-internal	simple textual		clause-final
	Rheme	Theme		Rheme

but	it	is often	difficult	to be	confident
textual initial	topical non-initial		clause-internal		clause-final
Theme			Rheme		

that	a given set	of security mechanisms
	--t specified unmarked	--t specified - asserting
	no-referent---addition	single preceding---anaphora complete-repetition
textual initial	topical non-initial	topical non-initial
Theme		

fully implements	the desired security policies.	In Section 2.3.3,
	--t specified - asserting	--t specified unmarked
	single preceding---anaphora complete-repetition	single non-verbal---exophora
	clause-complex-final	experiential initial
	Rheme	Theme

we	introduced	a security model
--t specified interlocuters		--t specified unmarked
context-of-culture---homophora single		no-referent---addition
topical non-initial		clause-final
		Rheme

that	is designed to help in analysing	the potential security threats
		--t specified - asserting
		single within-group---esphora meronymy-constitution facet
simple topical		clause-internal
Theme		Rheme

in a distributed system.	We	can summarize
--t specified unmarked	--t specified interlocuters	
no-referent---addition	single context-of-culture---homophora	
clause-complex-final	simple topical	
	Theme	

the security model	of Chapter 2	as follows:
--t specified - asserting	--t specified unrestricted	
single preceding---anaphora complete-repetition	single non-verbal---exophora	
clause-internal	clause-internal	clause-final
Rheme		

- Processes	encapsulate	resources	(such as
--t specified unmarked		--t specified unmarked	
no-referent---addition		no-referent---addition	
simple topical		clause-internal	
Theme		Rheme	

programming language- level objects	and other system-defined resources)	and
--t specified unmarked	specified general-comparison semblance-similarity non-interlocuters	
single preceding---anaphora elaboration	single preceding---anaphora attitudinal	
clause-internal	clause-final	textual simple
		Theme

allow	clients	to access	them
	--t specified unmarked		--t specified non-interlocuters
	no-referent---addition		preceding---anaphora multiple---ambiguous substitution
	clause-internal		clause-internal
	Rheme		

through their interfaces.	Principals	(users
--t specified - asserting	--t specified unmarked	--t specified unmarked
single preceding---anaphora alienable meronymy-relational	no-referent---addition	single preceding---anaphora complete-repetition
clause-complex-final	initial topical	topical non-initial
	Theme	

or other processes)	can be explicitly authorized to operate	on resources.
specified general-comparison difference - asserting		--t generic unmarked
single preceding---anaphora attitudinal		no-referent---addition
topical non-initial		clause-complex-final
		Rheme

Resources	must be protected against	unauthorized access.
generic --t - asserting		--t specified unmarked
single preceding---anaphora complete-repetition		multiple---ambiguous preceding---anaphora derivation
simple topical		clause-complex-final
Theme		Rheme

- Processes	interact through	a network	that
--t specified unmarked		--t specified unmarked	

no-referent---addition		no-referent---addition	
simple topical		clause-final	simple topical
Theme		Rheme	Theme

is shared by	many users.	Enemies
	--t specified particular	--t specified unmarked
	single preceding---anaphora inflexion	no-referent---addition
	clause-complex-final	topical initial
	Rheme	Theme

(attackers)	can access	the network.
--t specified unmarked		--t specified - asserting
single preceding---anaphora experiential		single preceding---anaphora complete-repetition
non-initial topical		clause-complex-final
Theme		Rheme

They	can copy or attempt to read	any message
--t specified non-interlocuters		--t specified unrestricted-2
single preceding---anaphora substitution		no-referent---addition
simple topical		clause-internal
Theme		Rheme

transmitted	through the network	and
	--t specified - asserting	
	single preceding---anaphora complete-repetition	
	clause-final	textual initial
	Rheme	Theme

they	can inject	arbitrary messages,
--t specified non-interlocuters		--t specified unmarked
single preceding---anaphora substitution		no-referent---addition
topical non-initial		clause-internal
		Rheme

addressed to	any destination	and purporting to come from
	--t specified unrestricted-2	
	no-referent---addition	
	clause-internal	
	Rheme	

any source,	into the network.	That security model
--t specified unrestricted-2	--t specified - asserting	--t specified distant
no-referent---addition	single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition
clause-internal	clause-complex-final	simple topical
		Theme

identifies	the features	of distributed systems
	--t specified - asserting	--t generic - asserting
	single within-group---esphora facet meronymy-constitution	single preceding---anaphora complete-repetition
	clause-internal	clause-final
	Rheme	

that	expose	them
		--t specified non-interlocuters
		multiple---ambiguous preceding---anaphora substitution
simple topical		clause-internal
Theme		Rheme

to attacks.	In this chapter,	we
--t generic - asserting	--t specified proximate	--t specified interlocuters
single preceding---anaphora complete-repetition	single non-verbal---exophora	context-of-culture---homophora single
clause-complex-final	experiential initial	topical non-initial
	Theme	

shall detail	these attacks	and	the security techniques
	--t specified proximate		--t specified - asserting
	single preceding---anaphora complete-repetition		single preceding---anaphora complete-repetition
	clause-final		clause-final
	Rheme		

that	are available for defeating	them.
		specified --t non-interlocuters
		single preceding---anaphora substitution
simple topical		clause-complex-final
Theme		Rheme

Appendix 3.5 Combination of Participant Identification and Tracking, Theme and Information Analyses for "Retardation Methods" Text

RETARDATION METHODS	INTRODUCTION	1.	For many years
neutralised	neutralised		specified --t non-particular
no-referent---addition	no-referent---addition		context-of-culture---homophora single
			initial experiential
			Theme

fixed wing aircraft	relied purely on	wheel brakes	to retard
--t generic unmarked		--t specified unmarked	
no-referent---addition		no-referent---addition	
topical non-initial		clause-internal	
		Rheme	

their forward movement	after	landing.	Over the years,
--t specified - asserting		specified --t non-particular	--t specified - asserting
single preceding---anaphora facet meronymy-constitution		context-of-culture---homophora single	context-of-culture---homophora single
clause-internal		clause-complex-final	initial experiential
Rheme			Theme

advancements	in design	and the introduction
--t specified unmarked	--t specified unmarked	--t specified - asserting
single within-group---esphora extension	no-referent---addition	single within-group---esphora extension
non-initial topical	non-initial topical	non-initial topical

of new materials	has led to	increased capabilities
--t specified unmarked		--t specified unmarked
no-referent---addition		no-referent---addition
topical non-initial		clause-internal
		Rheme

and carrying capacities,	which	has inevitably led to	faster and heavier aircraft.
--t specified unmarked			specified difference -- comparison-quality - asserting
no-referent---addition			context-of-culture---homophora single
clause-final	simple topical		clause-complex-final
	Theme		Rheme

Increased flying speeds	have led	in turn to

specified comparison-quality -- difference asserting -		
single preceding---anaphora experiential		
simple topical		
Theme		

higher landing speeds,	which	has increased
specified comparison-quality -- difference asserting -		
single preceding---anaphora experiential		
clause-final	simple topical	
Rheme	Theme	

the burden	placed on	the wheel brakes.
specified --t asserting -		specified --t asserting -
single within-group---esphora alienable meronymy-constitution		single preceding---anaphora complete-repetition
clause-internal		clause-complex-final
Rheme		Rheme

2.	Without some form	of additional assistance
	generic --t non-particular	specified --t unmarked
	single within-group---esphora relational hyponymy---class-preceding	no-referent---addition
	experiential initial	experiential initial
	Theme	

in retarding	the aircraft,	the wheel brakes
	--t specified - asserting	specified --t asserting -
	single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition
experiential initial	experiential initial	non-initial topical

would need to be applied in	a heavy and sustained manner	leading to
	specified --t unmarked	
	no-referent---addition	
	clause-final	
	Rheme	

rapid wear	of the brakes	and tyres,
specified --t unmarked	specified --t asserting -	specified --t asserting -
no-referent---addition	single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition
clause-internal	clause-internal	clause-final
Rheme		

or	runways	would need to be extended
	generic --t asserting -	
	context-of-culture---homophora single	

initial textual	non-initial topical	
Theme		

to unacceptable lengths.	To help reduce	brake
specified --t unmarked		generic --t asserting -
no-referent---addition		single preceding---anaphora complete-repetition
clause-complex-final	initial experiential	initial experiential
Rheme	Theme	

and tyre wear	and	to maintain
generic --t asserting -		
single preceding---anaphora complete-repetition		
initial experiential	non-initial textual	non-initial experiential

runways	at an acceptable length,	additional retardation methods
generic --t asserting -	specified --t unmarked	specified --t particular
single preceding---anaphora complete-repetition	no-referent---addition	single preceding---anaphora inflexion
experiential non-initial	experiential non-initial	non-initial topical

based on	the principle	of aerodynamic braking
	--t specified - asserting	--t specified unmarked
	single within-group---esphora relational hyponymy---class-preceding	no-referent---addition
	non-initial topical	non-initial topical

have been devised to supplement	the wheel brakes.	3.
	--t specified - asserting	
	single preceding---anaphora complete-repetition	
	clause-complex-final	
	Rheme	

Methods	of emergency retardation	have also been devised to arrest
--t generic - asserting	--t specified unmarked	
single preceding---anaphora complete-repetition	single preceding---anaphora attitudinal	
simple topical	simple topical	
Theme		

the forward motion	of the aircraft	on the ground
--t specified - asserting	--t specified - asserting	--t specified - asserting
single preceding---anaphora experiential	single preceding---anaphora complete-repetition	context-of-culture---homophora single

clause-internal	clause-internal	clause-internal
Rheme		

in the event	of failure	of the wheel brake system
--t specified - asserting	--t specified unmarked	--t specified - asserting
single within-group---esphora relational hyponymy---class-preceding	no-referent---addition	single preceding---anaphora experiential
clause-internal	clause-final	clause-final

or	during an aborted take-off	where	normal braking methods
	--t specified unmarked		--t specified unmarked
	no-referent---addition		single preceding---anaphora attitudinal
clause-internal	clause-final	textual initial	topical non-initial
		Theme	

would be	inadequate.	WHEEL BRAKES	4.	Wheel brakes
		neutralised		generic --t asserting -
		no-referent---addition		no-referent---addition
	clause-complex-final			simple topical
	Rheme			Theme

form	the primary method	of retarding
	--t specified - asserting	generic --t - asserting
	single preceding---anaphora relational hyponymy---class-preceding	single preceding---anaphora complete-repetition
	clause-internal	clause-internal
	Rheme	

the forward movement	of most aircraft	when
--t specified - asserting	specified --t particular	
single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	
clause-internal	clause-internal	clause-internal

on the ground	and,	in common with
--t specified - asserting		
single preceding---anaphora complete-repetition		
clause-final	textual initial	
	Theme	

most braking systems,	they	rely on
--t specified particular	--t specified - asserting	
single preceding---anaphora attitudinal	single preceding---anaphora substitution	
experiential non-initial	topical non-initial	

the principle	of energy conversion	for their operation.
--t specified - asserting	generic --t - asserting	--t specified - asserting
single within-group---esphora relational hyponymy---class-preceding	no-referent---addition	single preceding---anaphora extension
clause-internal	clause-internal	clause-complex-final
Rheme		

In this method,	friction	is used to convert
--t specified proximate	--t specified unmarked	
single preceding---anaphora attitudinal	no-referent---addition	
experiential initial	topical non-initial	
Theme		

the forward motion	of the aircraft	(kinetic energy)
--t specified - asserting	--t specified - asserting	--t specified unmarked
single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	no-referent---addition
clause-internal	clause-internal	clause-internal
Rheme		

into heat,	which	is then dissipated,	to the surrounding air.
--t generic - asserting	--t specified non-interlocuters		--t generic - asserting
no-referent---addition	single preceding---anaphora substitution		context-of-culture---homophora single
clause-final	simple topical		clause-complex-final
	Theme		Rheme

5.	Wheel brakes	and their associated operating systems
	--t generic - asserting	--t specified - asserting
	single preceding---anaphora complete-repetition	single preceding---anaphora meronymy-constitution facet
	simple topical	simple topical
	Theme	

are covered	in Section 3 Chapters 1 and 3	of this Volume.
	--t specified unmarked	--t specified proximate
	single non-verbal---exophora	single non-verbal---exophora
	clause-internal	clause-complex-final
	Rheme	

AERODYNAMIC BRAKING	6.	The term	aerodynamic braking
neutralised		--t asserting - generic	--t specified unmarked
no-referent---addition		context-of-culture---homophora single	no-referent---addition
		simple topical	Simple topical
		Theme	

means	using	the airflow	to assist in
		generic --t asserting -	
		context-of-culture---homophora single	
		clause-internal	
		Rheme	

retardation	by selectively increasing	the amount
--t specified - asserting		--t specified - asserting
single preceding---anaphora complete-repetition		single within-group---esphora relational hyponymy---class-preceding
clause-final		clause-internal

of drag	produced by	the aircraft.
--t specified unmarked		--t specified - asserting
no-referent---addition		single preceding---anaphora complete-repetition
clause-internal		clause-complex-final

This increase in	drag	is normally applied,
--t specified proximate	--t specified - asserting	
single preceding---anaphora derivation	single preceding---anaphora complete-repetition	
simple topical	simple topical	
Theme		

in conjunction with	operation of	the wheel brakes,
	generic --t - asserting	--t specified - asserting
	single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition
	clause-internal	clause-final
	Rheme	

immediately after	landing	or
	--t generic - asserting	
	single preceding---anaphora complete-repetition	
clause-internal	clause-internal	clause-internal

during	an aborted take-off	and	is produced by
	--t specified unmarked		
	no-referent---addition		
clause-internal	clause-final	textual simple	
		Theme	

one,	or sometimes	a combination of
--t specified non-interlocuters		--t specified unmarked
no-referent---addition		no-referent---addition

clause-internal	clause-internal	clause-internal
Rheme		

more than one	, of the following methods:	• Flying control surfaces.
specified difference -- comparison-quantity non-interlocuters	--t specified proximate	--t specified unmarked
no-referent---addition	single beyond-group---cataphora relational hyponymy---class-preceding	no-referent---addition
clause-internal	clause-complex-final	

• Thrust reversal.	• Brake parachute.
--t specified unmarked	--t specified unmarked
no-referent---addition	no-referent---addition

• Angle of Attack.	Flying control surfaces	7	Flying control surfaces
--t specified unmarked	neutralised		specified --t asserting -
no-referent---addition	no-referent---addition		single preceding---anaphora complete-repetition
			simple topical
			Theme

which	can be used to help retard	the aircraft
		--t generic - asserting
		single preceding---anaphora complete-repetition
simple topical		clause-internal
		Rheme

through aerodynamic braking	are:	• Airbrakes.
--t specified unmarked		--t specified unmarked
single preceding---anaphora complete-repetition		no-referent---addition
clause-internal	clause-final	

• Elevators	• Flaps.	• Tailerons	• Spoilers.
--t specified unmarked	--t specified unmarked	--t specified unmarked	--t specified unmarked
no-referent---addition	no-referent---addition	no-referent---addition	no-referent---addition

• Foreplanes	8	As	the airbrakes,
--t specified unmarked			--t specified - asserting
no-referent---addition			single preceding---anaphora complete-repetition

		textual initial	topical non-initial
		Theme	

flaps	or spoilers	are moved out
--t specified - asserting	--t specified - asserting	
single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	
topical non-initial	topical non-initial	

into the airflow	they	cause
--t generic - asserting	--t specified non-interlocuters	
single preceding---anaphora complete-repetition	single preceding---anaphora substitution	
clause-final	simple topical	
Rheme	Theme	

an increase in	drag.	The further
--t specified unmarked	--t generic - asserting	specified difference -- comparison-quality - asserting
no-referent---addition	single preceding---anaphora complete-repetition	single within-group---esphora alienable meronymy-relational
clause-internal	clause-complex-final	experiential initial
Rheme		Theme

into the airflow	they	protrude,
--t specified - asserting	--t specified non-interlocuters	
single preceding---anaphora complete-repetition	single preceding---anaphora substitution	
experiential non-initial	topical non-initial	

the greater	the drag	produced
specified difference -- comparison-quality - asserting	--t specified - asserting	
single within-group---esphora alienable meronymy-relational	single preceding---anaphora complete-repetition	
experiential initial	topical non-initial	
Theme		

and,	just as	in flight,	any increase in
		generic --t asserting -	--t specified unmarked
		single preceding---anaphora derivation	no-referent---addition
textual initial		experiential non-initial	topical non-initial

drag	slows down	the forward progress
--t generic - asserting		--t specified - asserting
single preceding---anaphora complete-repetition		single preceding---anaphora attitudinal
topical non-initial		clause-internal
		Rheme

of the aircraft.	It	should be noted that
--t specified - asserting		
single preceding---anaphora complete-repetition		
clause-complex-final	simple topical	
	Theme	

only the airbrake	is specially designed to slow	the aircraft;
specified --t asserting -		--t specified - asserting
single preceding---anaphora complete-repetition		single preceding---anaphora complete-repetition
simple topical		clause-final
		Rheme

for the other controls	this	is
specified general-comparison difference - asserting	--t specified proximate	
single preceding---anaphora superordination-class-subclass hyperonymy---subclass-preceding	multiple---ambiguous preceding---anaphora substitution	
experiential initial	topical non-initial	
Theme		

only a secondary function.	The flying control surfaces	which
--t specified unmarked	--t specified – asserting	--t specified non-interlocuters
no-referent---addition	single preceding---anaphora complete-repetition	single preceding---anaphora substitution
clause-complex-final	simple topical	simple topical
Rheme	Theme	Theme

provide	aerodynamic braking	are shown	at Fig 1.
	--t specified unmarked		--t specified unmarked
	single preceding---anaphora complete-repetition		single non-verbal---exophora
	clause-final		clause-complex-final
	Rheme		Rheme

Thrust reversal	9.	Thrust reversal	is used to assist in
neutralised		generic --t asserting -	
no-referent---addition		single preceding---anaphora complete-repetition	
		simple topical	
		Theme	

aircraft retardation	by using	engine power
generic --t asserting -		generic --t asserting -
single preceding---anaphora experiential		no-referent---addition
clause-internal		clause-internal
Rheme		

as a deceleration force.	10.	Propeller driven aircraft.	On propeller driven aircraft,

specified --t unmarked		neutralised	--t generic - asserting
no-referent---addition		no-referent---addition	single preceding---anaphora complete-repetition
clause-complex-final			experiential initial
			Theme

thrust reversal	is achieved by reversing	the pitch
--t specified unmarked		--t specified - asserting
single preceding---anaphora complete-repetition		single within-group---esphora facet meronymy-constitution
topical non-initial		clause-internal
		Rheme

of the propeller blades	so that	the thrust generated
--t specified - asserting		--t specified - asserting
single preceding---anaphora meronymy-constitution part		single preceding---anaphora derivation
clause-final	textual initial	topical non-initial
	Theme	

by the propeller	is directed	forwards	instead of
--t specified - asserting			
single preceding---anaphora derivation			
topical non-initial		clause-internal	clause-internal
		Rheme	

rearwards,	the degree	of braking assistance
	--t specified - asserting	--t specified unmarked
	single within-group---esphora measure meronymy-relational	no-referent---addition
clause-final	simple topical	simple topical
	Theme	

being controlled by	use of	the engine throttle lever(s).
	--t generic - asserting	specified --t - asserting
	single within-group---esphora extension	single preceding---anaphora meronymy-constitution part
	clause-internal	clause-complex-final
	Rheme	

11.	Turbo-jet aircraft.	Thrust reversal	on turbo-jet aircraft
	neutralised	--t specified unmarked	specified --t unmarked
	no-referent---addition	single preceding---anaphora complete-repetition	no-referent---addition
		simple topical	simple topical
		Theme	

is achieved by changing	the direction	of the exhaust gas stream
	--t specified - asserting	--t specified - asserting

	single within-group---esphora facet meronymy-relational	context-of-culture---homophora single
	clause-internal	clause-internal
	Rheme	

through more than 90 degrees,	but	less than 180 degrees.
specified comparison-quantity -- difference asserting -		specified difference -- comparison-quantity - asserting
single preceding---anaphora measure meronymy-constitution		single preceding---anaphora meronymy-constitution measure
clause-internal	clause-internal	clause-complex-final

One method	of deflecting	the exhaust gas stream
--t specified particular	--t specified unmarked	--t specified - asserting
no-referent---addition	no-referent---addition	single preceding---anaphora complete-repetition
simple topical	simple topical	simple topical
Theme		

is shown	at Fig 2.	During flight
	--t specified unmarked	generic --t asserting -
	single non-verbal---exophora	single preceding---anaphora complete-repetition
	clause-complex-final	experiential initial
	Rheme	Theme

the bucket-type doors	are held	in the open position
--t specified - asserting		--t specified - asserting
context-of-culture---homophora single		context-of-culture---homophora single
topical non-initial		clause-final
		Rheme

as shown	in Fig 2A,	allowing	rearward passage
	--t specified unmarked		--t specified unmarked
	single non-verbal---exophora		no-referent---addition
	clause-final		clause-internal

of the exhaust gas stream.	When	braking assistance
--t specified - asserting		--t generic - asserting
single preceding---anaphora complete-repetition		single preceding---anaphora complete-repetition
clause-complex-final	textual initial	topical non-initial
	Theme	

is	required,	the doors	are moved into
		--t specified - asserting	
		single preceding---anaphora derivation	
	clause-final	simple topical	
	Rheme	Theme	

the gas stream	by hydraulic	or pneumatic jacks
--t specified - asserting	--t specified unmarked	--t specified unmarked
single preceding---anaphora derivation	no-referent---addition	no-referent---addition
clause-internal	clause-internal	clause-final
Rheme		

as shown	in Fig 2B,	uncovering	apertures
	--t specified unmarked		--t generic - asserting
	single non-verbal---exophora		no-referent---addition
	clause-final		clause-final

through which	the gas	is deflected
--t specified non-interlocuters	--t specified - asserting	
single preceding---anaphora substitution	single preceding---anaphora derivation	
textual initial	topical non-initial	
Theme		

in a slightly forward direction	to give	'thrust reversal'.	Brake parachute	12.
--t specified unmarked		specified --t unique--i	neutralised	
no-referent---addition		no-referent---addition	no-referent---addition	
clause-internal		clause-complex-final		
Rheme				

Brake parachutes	are normally to be found on	fast jet aircraft
specified --t unmarked		specified --t unmarked
single preceding---anaphora complete-repetition		no-referent---addition
simple topical		clause-final
Theme		Rheme

and	produce	aerodynamic drag	to assist	the wheel brakes
		specified --t unmarked		--t specified - asserting
		no-referent---addition		single preceding---anaphora complete-repetition
textual simple		clause-internal		clause-internal
Theme		Rheme		

in retarding	the aircraft	during its landing run.
	--t specified - asserting	--t specified - asserting
	single preceding---anaphora complete-repetition	single preceding---anaphora extension
clause-internal	clause-internal	clause-complex-final

A typical brake parachute installation	consists of	the parachute assembly
--t specified unmarked		--t specified - asserting
no-referent---addition		single preceding---anaphora attitudinal
simple topical		clause-internal
Theme		Rheme

in housing	with associated release/ jettison controls and mechanisms	(Fig 3).
--t specified unmarked	--t specified unmarked	--t specified unmarked
no-referent---addition	no-referent---addition	single non-verbal---exophora
clause-internal	clause-internal	clause-complex-final

13.	The parachute assembly.	The parachute assembly
	neutralised	--t specified - asserting
	no-referent---addition	single preceding---anaphora complete-repetition
		simple topical
		Theme

consists of	the main parachute,	the auxiliary (or drogue) parachute
	--t specified - asserting	--t specified - asserting
	single preceding---anaphora meronymy-constitution part	single preceding---anaphora meronymy-constitution part
	clause-internal	clause-internal
	Rheme	

and the streamer cable.	A ring	on the free end

--t specified - asserting	--t specified unmarked	--t specified - asserting
single preceding---anaphora meronymy-constitution part	no-referent---addition	single within-group---esphora meronymy-constitution part
clause-complex-final	simple topical	simple topical
	Theme	

of the streamer cable	connects	the parachute assembly
--t specified - asserting		--t specified - asserting
single preceding---anaphora complete-repetition		single preceding---anaphora complete-repetition
simple topical		clause-internal
		Rheme

to the airframe	through a release unit.	The drogue
--t specified - asserting	--t specified unmarked	--t specified - asserting
single preceding---anaphora meronymy-constitution part	no-referent---addition	single preceding---anaphora derivation
clause-internal	clause-complex-final	simple topical
		Theme

may be	a small, spring-loaded version	of the main parachute	or
	--t specified unmarked	--t specified - asserting	
	no-referent---addition	single preceding---anaphora complete-repetition	
	clause-internal	clause-final	textual simple
	Rheme		

connected to	a rigid conical cap,	which	forms
	--t specified unmarked	--t specified non-interlocuters	
	no-referent---addition	single preceding---anaphora substitution	
	clause-final	textual simple	
		Theme	

the tip of	the a/c tail cone	during flight.
--t specified - asserting	--t specified – asserting	--t generic - asserting
single within-group---esphora part meronymy-constitution	context-of-culture---homophora single	single preceding---anaphora complete-repetition
clause-internal	clause-internal	clause-complex-final
Rheme		

The main parachute	and streamer cable	are packed into
--t specified - asserting	--t specified unmarked	
single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	
simple topical	simple topical	
Theme		

a canvas bag	or light alloy canister	(hopper),	whilst
--t specified unmarked	--t specified unmarked	--t specified unmarked	

no-referent---addition	no-referent---addition	no-referent---addition	
clause-internal	clause-internal	clause-final	textual initial
Rheme			Theme

the drogue	may be packed with	the main parachute
--t specified - asserting		--t specified - asserting
single preceding---anaphora complete-repetition		single preceding---anaphora complete-repetition
topical non-initial		clause-internal
		Rheme

and streamer cable	or	may be located	separately.
--t specified - asserting			
single preceding---anaphora complete-repetition			
clause-final	textual simple		clause-complex-final
	Theme		Rheme

The main parachute canopy	is made of	vented nylon panels
--t specified - asserting		--t specified unmarked
single preceding---anaphora meronymy-constitution part		no-referent---addition
simple topical		clause-internal
Theme		Rheme

or ribbons,	connected to	the streamer cable
--t specified unmarked		--t specified - asserting
no-referent---addition		single preceding---anaphora complete-repetition
clause-final		clause-internal

by nylon rigging lines	and a connector block.	14. The parachute housing.
--t specified unmarked	--t specified unmarked	neutralised
no-referent---addition	no-referent---addition	no-referent---addition
clause-internal	clause-complex-final	

A typical brake parachute housing	is located	as near as possible to
--t specified unmarked		
single preceding---anaphora derivation		
simple topical		clause-internal
Theme		Rheme

the end of	the rear fuselage.	The housing
--t specified - asserting	--t specified - asserting	--t specified - asserting
single within-group---esphora meronymy-constitution part	context-of-culture---homophora single	single preceding---anaphora derivation
clause-internal	clause-complex-final	simple topical
		Theme

comprises of	a compartment	designed to accommodate	the parachute assembly
	--t specified unmarked		--t specified - asserting
	no-referent---addition		single preceding---anaphora complete-repetition
	clause-internal		clause-internal
	Rheme		

or hopper	with a spring loaded door	or detachable cap,
--t specified - asserting	--t specified unmarked	--t specified unmarked
single preceding---anaphora complete-repetition	no-referent---addition	no-referent---addition
clause-internal	clause-internal	clause-final

which	closes off	the opening
--t specified non-interlocuters		--t specified - asserting
single preceding---anaphora substitution		single part meronymy-constitution within-group---esphora
textual simple		clause-internal
Theme		Rheme

to the compartment.	15.	The release and jettison controls and mechanisms.
--t specified - asserting		neutralised
single preceding---anaphora complete-repetition		no-referent---addition
clause-complex-final		

Control of	the brake parachute	is normally effected mechanically by
--t generic - asserting	--t specified - asserting	
single within-group---esphora extension	single preceding---anaphora complete-repetition	
simple topical	simple topical	
Theme		

a Teleflex	or Bowden control cable system,	but	it
--t specified unmarked	--t specified unmarked		--t specified non-interlocuters
no-referent---addition	no-referent---addition		single preceding---anaphora substitution
clause-internal	clause-final	textual initial	topical non-initial
Rheme		Theme	

may be achieved electrically or,	on some aircraft,	by a combination
	specified --t non-particular	--t specified unmarked
	no-referent---addition	no-referent---addition
	clause-internal	clause-internal
	Rheme	

of both methods.	Cockpit selection	for both the release
specified --t asserting inclusive	--t specified unmarked	--t specified asserting inclusive
single preceding---anaphora attitudinal	single within-group---esphora enhancement	single within-group---esphora extension
clause-complex-final	simple topical	simple topical
	Theme	

and jettison of	the brake parachute	is usually made from
--t specified asserting inclusive	--t specified - asserting	
single extension within-group---esphora	single preceding---anaphora complete-repetition	
simple topical	simple topical	

a single control,	except for	combination systems	where
--t specified unmarked		--t specified unmarked	
no-referent---addition		no-referent---addition	
clause-final	clause-internal	clause-final	textual initial
Rheme			Theme

the parachute	is released mechanically by	a lever
--t specified - asserting		--t specified unmarked
single preceding---anaphora complete-repetition		no-referent---addition
topical non-initial		clause-internal
		Rheme

or handle	and	a separate switch	is used to operate
--t specified unmarked		--t specified unmarked	
no-referent---addition		no-referent---addition	
clause-final	textual initial	topical non-initial	
	Theme		

an electrical jettison unit.	The cockpit control	is normally guarded to prevent
--t specified unmarked	--t specified - asserting	
no-referent---addition	single preceding---anaphora derivation	
clause-complex-final	simple topical	
Rheme	Theme	

inadvertent operation.	16.	Principles of operation	When
--t specified unmarked		neutralised	
no-referent---addition		no-referent---addition	
clause-complex-final			textual initial
Rheme			Theme

'release'	or 'stream'	is selected at
--t specified particular	specified --t particular	
single preceding---anaphora derivation	single preceding---anaphora derivation	

topical non-initial	topical non-initial	

the cockpit control,	the door	or end cap
--t specified - asserting	--t specified - asserting	--t specified - asserting
single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition
clause-final	simple topical	simple topical
Rheme	Theme	

of the compartment housing	the brake parachute	opens to release
--t specified - asserting	--t specified - asserting	
single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	
simple topical	simple topical	

the drogue	into the airstream	as shown in
--t specified - asserting	--t specified - asserting	
single preceding---anaphora complete-repetition	context-of-culture---homophora single	
clause-internal	clause-final	
Rheme		

Fig 4A.	17.	The drag	of the drogue parachute
--t specified unmarked		--t specified - asserting	--t specified - asserting
single non-verbal---exophora		single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition
clause-complex-final		simple topical	simple topical
		Theme	

withdraws	the main parachute	from its pack
	--t specified - asserting	--t specified - asserting
	single preceding---anaphora complete-repetition	single preceding---anaphora meronymy-constitution part
	clause-internal	clause-internal
	Rheme	

along with a streamer cable,	as	the cable
--t specified unmarked		--t specified - asserting
no-referent---addition		single preceding---anaphora derivation
clause-final	textual initial	topical non-initial
	Theme	

is pulled	taut	the main parachute	develops (opens) to provide
		--t specified - asserting	
		single preceding---anaphora complete-repetition	
	clause-final	simple topical	
	Rheme	Theme	

the braking drag	as shown	in Fig 4B.
--t specified - asserting		--t specified unmarked
single preceding---anaphora derivation		single non-verbal---exophora
clause-final		clause-complex-final
Rheme		

NOTES	(1)	On brake parachute installations	fitted with a door,
		--t specified unmarked	--t specified unmarked
		no-referent---addition	no-referent---addition
		experiential initial	experiential initial
		Theme	

the door	is usually designed to hinge open	under the influence of
--t specified - asserting		--t specified - asserting
single preceding---anaphora complete-repetition		single within-group---esphora extension
topical non-initial		clause-internal
		Rheme

a damped spring	when	the parachute release control
--t specified unmarked		--t specified - asserting
no-referent---addition		single preceding---anaphora derivation
clause-final	initial textual	simple topical
	Theme	

is operated,	allowing	the spring loaded drogue parachute	to deploy
		--t specified - asserting	
		single preceding---anaphora derivation	
		clause-final	
		Rheme	

and	withdraw	the main parachute.
		--t specified - asserting
		single preceding---anaphora complete-repetition
simple textual		clause-complex-final
Theme		Rheme

(2)	On those installations	where	the parachute housing
	--t specified distant		--t specified - asserting
	single preceding---anaphora substitution		single preceding---anaphora derivation
	experiential initial	textual initial	topical non-initial
	Theme	Theme	

is closed off by	a cap,	the cap
	--t specified unmarked	--t specified - asserting
	no-referent---addition	single preceding---anaphora complete-repetition
	clause-final	topical non-initial

	Rheme	Theme

is connected to	the drogue chute	by nylon cords.
	--t specified - asserting	--t specified unmarked
	single preceding---anaphora complete-repetition	no-referent---addition
	clause-internal	clause-complex-final
	Rheme	

When	the parachute release control	is operated
	--t specified - asserting	
	single preceding---anaphora complete-repetition	
textual initial	topical non-initial	
Theme		

on this type of installation,	the cap	is completely released from
--t specified proximate	--t specified - asserting	
single preceding---anaphora complete-repetition	single preceding---anaphora complete-repetition	
clause-final	simple topical	
Rheme	Theme	

the airframe structure	and,	as	it
--t specified - asserting			--t specified non-interlocuters
single preceding---anaphora derivation			single preceding---anaphora substitution
clause-final	textual initial	textual initial	topical non-initial
Rheme	Theme	Theme	

is carried away by	the airflow,	it
	--t specified - asserting	--t specified non-interlocuters
	single preceding---anaphora complete-repetition	single preceding---anaphora substitution
	clause-internal	topical non-initial
	Rheme	Theme

withdraws	the drogue chute	which
	--t specified - asserting	--t specified non-interlocuters
	single preceding---anaphora complete-repetition	single preceding---anaphora substitution
	clause-final	textual simple
	Rheme	Theme

withdraws	the main parachute.	18.	Once
	--t specified - asserting		
	single preceding---anaphora complete-repetition		
	clause-complex-final		textual initial
	Rheme		Theme

the aircraft	has slowed to	a speed
--t specified - asserting		--t specified unmarked
single preceding---anaphora complete-repetition		no-referent---addition
topical non-initial		clause-final
		Rheme

where	the brake parachute	is	is no longer effective,
	--t specified - asserting		
	single preceding---anaphora complete-repetition		
textual initial	non-initial topical		clause-final
Theme			Rheme

it	is jettisoned	onto the runway
--t specified non-interlocuters		generic --t - asserting
single preceding---anaphora substitution		single preceding---anaphora complete-repetition
simple topical		clause-internal
Theme		Rheme

or taxiway	to be collected by	a brake parachute recovery team.
--t generic - asserting		--t specified unmarked
single preceding---anaphora experiential		no-referent---addition
clause-internal		clause-complex-final

Angle of attack.	19.	On some aircraft types,	during landing,
Neutralised		--t specified particular	--t generic - asserting

no-referent---addition		no-referent---addition	single preceding---anaphora complete-repetition
		experiential initial	experiential non-initial
		Theme	

the pilot	of the aircraft	will hold
--t specified - asserting	--t specified - asserting	
single within-group---esphora meronymy-constitution part	single preceding---anaphora complete-repetition	
topical non-initial	topical non-initial	

the nose	of the aircraft	high
--t specified - asserting	--t specified - asserting	
single preceding---anaphora meronymy-constitution part	single preceding---anaphora complete-repetition	
clause-internal	clause-internal	clause-internal
Rheme		

off the runway,	rolling along on	his main wheels
--t specified - asserting		--t specified - asserting
single preceding---anaphora complete-repetition		single preceding---anaphora meronymy-constitution part
clause-final		clause-internal

only.	This high angle of attack	produces	a lot of aerodynamic drag,
	--t specified proximate		specified --t unmarked
	single preceding---anaphora attitudinal		no-referent---addition
clause-complex-final	simple topical		clause-final
	Theme		Rheme

which	assists	the retardation
--t specified non-interlocuters		--t specified - asserting
single preceding---anaphora substitution		single within-group---esphora extension
textual simple		clause-internal
Theme		Rheme

of the aircraft.
--t specified - asserting
single preceding---anaphora complete-repetition
clause-complex-final

Appendix 4.1 χ^2 Contingency Tables

Note 1:
The following method has been used to derive the value for χ^2

	1	2			χ^2 table
A	*O= Observed Score*	*O= Observed Score*	Total Row A	A1	**$(E-O)^2/E$**
E = Expected Scores	**E = (Row A * Column 1) / Grand Total**	**E = (Row A * Column2) / Grand Total**		A2	**$(E-O)^2/E$**
B	*O= Observed Score*	*O= Observed Score*	Total Row B	B1	**$(E-O)^2/E$**
E = Expected Scores	**E = (Row B * Column1) / Grand Total**	**E = (Row B * Column2) / Grand Total**		B2	**$(E-O)^2/E$**
	Total Column 1	Total Column 2	Grand Total	χ^2 =	**Sum A1-B2 $\Sigma (E-O)^2/E$**
				phi=	**$\sqrt{\chi^2}$/ Grand Total**

Note 2:
The following method has been used to derive the Percentage Deviations for each cell

	1	2
A	**(Observed-Expected)/Expected *100**	**(Observed-Expected)/Expected *100**
B	**(Observed-Expected)/Expected *100**	**(Observed-Expected)/Expected *100**

Table 1 Contingency Table comparing Participant/Non-Participant Groups with position in Theme/Information

	Participant	Non- Participant			χ^2 table
Information	*1246*	*253*	1499		**18.23474**
E =	**1104.109**	**394.8913**			**28.56205**
Theme	*563*	*394*	957		**50.984**
E =	**704.8913**	**252.1087**			**79.85895**
	1809	647	2456	critical value=	**177.6397**
				phi=	**0.26894**

Table 2 Contingency Table comparing Participant/Non-Participant Groups with informational position

	Participant	Non- Participant			χ^2 table
Clause-internal	*680*	*111*	791		**0.770264**
E =	**657.4957**	**133.5043**			**0.860563**
Clause Final	*566*	*142*	708		**3.793473**
E =	**588.5043**	**119.4957**			**4.238189**
	1246	253	1499	critical value=	**9.662489**
				phi=	**0.080287**

Table 3 Contingency Table comparing Initial/Final Clausal Position with Participant Identification status

	Simple or Initial Theme	Clause / Complex Final			χ^2 table
Participant	*579*	*634*	1213		**9.743549**
E =	**659.1396**	**553.8604**			**31.34993**
Non-Participant	*285*	*92*	377		**11.59563**
E =	**204.8604**	**172.1396**			**37.30901**
	864	726	1590	χ2=	**89.99812**
				phi=	**0.237913**

Table 4 Contingency Table comparing Initial/Final Clausal Position with Reference type

	Simple or Initial Theme	Clause / Complex Final			χ^2 table
Presuming	*366*	*310*	676		**5.938838**
E =	**322.2529**	**353.7471**			**7.632423**
Presenting	*207*	*319*	526		**5.410102**
E =	**250.7471**	**275.2529**			**6.952906**
	573	629	1202	χ2=	**25.93427**
				phi=	**0.146887**

Table 5 Contingency Table comparing Initial/Final Clausal Position with Presumed Nominal/Pronominal participant

a. Contingency Table

	Simple or Initial Theme	Clause / Complex Final			χ^2 table
Presuming Nominal	*278*	*279*	557		**2.219193**
E =	**303.9726**	**253.0274**			**12.48577**
Presuming Pronominal	*80*	*19*	99		**2.666011**
E =	**54.02744**	**44.97256**			**14.99968**
	358	298	656	χ2=	**32.37065**
				phi=	**0.222138**

b. Percentage Deviation

	Simple or Initial Theme	Clause / Complex Final
Presuming Nominal	**-8.54438**	**10.26472**
Presuming Pronominal	**48.07291**	**-57.752**

Table 6 Contingency Table comparing Clausal Initial/Final Position with Presumed Pronominal participant

	Simple or Initial Theme	Clause / Complex Final			χ^2 table
Undirected	*247*	*268*	515		**2.712381**
E =	**274.2752**	**240.7248**			**14.10986**
Directed	*80*	*19*	99		**3.090413**
E =	**52.72476**	**46.27524**			**16.07639**
	327	287	614	χ2=	**35.98905**
				phi=	**0.242103**

Table 7 Contingency Table comparing Clausal Initial/Final Position with Presenting Marked/Unmarked

	Simple or Initial Theme	Clause / Complex Final			χ^2 table
Presenting Unmarked	*164*	*274*			
			438		**0.020715**
E =	**165.8536**	**272.1464**			**0.22683**
Presenting Marked	*17*	*23*			
			40		**0.012624**
E =	**15.14644**	**24.85356**			**0.138237**
	181	297	478	χ2=	**0.398406**
				phi=	**0.02887**

Table 8 Contingency Table comparing Clausal Initial/Final Position with Participant Tracking status

a. Contingency Table

	Theme	Information			χ^2 table
Participant	*902*	*1243*	2145		**7.62458**
E =	**988.8298**	**1156.17**			**30.62682**
Non-Participant	*333*	*201*	534		**6.521022**
E =	**246.1702**	**287.8298**			**26.19399**
	1235	1444	2679	χ2=	**70.96642**
				phi=	**0.162757**

b. Percentage Deviation

	Theme	Information
Participant	**-8.78107**	**7.510121**
Non- Participant	**35.27226**	**-30.1671**

Table 9 Contingency Table comparing Theme/Information clausal positions with Referent/Addition

	Theme	Information			χ^2 table
No referent – Addition	*224*	*515*	739		**24.22173**
E =	**310.759**	**428.241**			**12.73105**
Referent	*678*	*728*	1406		**17.57683**
E =	**591.241**	**814.759**			**9.238462**
	902	1243	2145	χ2=	**63.76807**
				phi=	**0.17242**

Table 10 Contingency Table comparing Theme/Information clausal positions with Tracked Context

	Theme	Information			χ^2 table
Context of Culture (hompohora)	*27*	*48*	75		**2.323244**
E =	**36.16643**	**38.83357**			**0.130912**
Context of situation	*651*	*680*	1331		**2.16368**
E =	**641.8336**	**689.1664**			**0.12192**
	678	728	1406	χ2=	**4.739756**
				phi=	**0.058061**

Table 11 Contingency Table comparing Theme/Information clausal positions with Lexical Relations

a. Contingency Table

	Theme	Information			χ^2 table
Superordination	*528*	*476*	1004	A1	**3.302772**
E =	**487.8592**	**516.1408**		A2	**3.121798**
Composition	*67*	*144*	211	B1	**12.31126**
E =	**102.5282**	**108.4718**		B2	**11.63667**
Nuclear Relations	*26*	*37*	63	C1	**0.695032**
E =	**30.61268**	**32.38732**		C2	**0.656948**
	621	657	1278	χ^2 =	**31.72448**
				phi=	**0.157555**

b. Percentage Deviation

	Theme	Information
Superordination	**8.227958**	**-7.77711**
Composition	**-34.6521**	**32.75336**
Nuclear Relations	**-15.0679**	**14.24223**

Table 12 Contingency Table comparing Theme/Information clausal positions with Superordination type

	Theme	Information			χ^2 table
Co-hyponymy	*492*	*422*	914		**0.267095**
E =	**480.6693**	**433.3307**			**2.712495**
Class-Subclass	*36*	*54*	90		**0.296273**
E =	**47.33068**	**42.66932**			**3.008819**
	528	476	1004	χ2=	**6.284682**
				phi=	**0.079118**

Table 13 Contingency Table comparing Theme/Information clausal positions with Repetition/Substitution

	Theme	Information			χ^2 table
Complete Repetition	*243*	*272*	515		**4.695457**
E =	**279.2078**	**235.7922**			**13.9778**
Substitution	*130*	*43*	173		**5.560017**
E =	**93.79215**	**79.20785**			**16.5515**
	373	315	688	χ2=	**40.78477**
				phi=	**0.243475**

Table 14 Contingency Table comparing Theme/Information clausal positions with Preceding/Following reference

a. Contingency Table

	Theme	Information			χ^2 table
Preceding – Anaphora	*520*	*511*	1031		**0.722193**
E =	**500.9789**	**530.0211**			**3.014496**
Following	*101*	*146*	247		**0.68262**
E =	**120.0211**	**126.9789**			**2.849319**
	621	657	1278	χ2=	**7.268628**
				phi=	**0.075416**

b. Percentage Deviation

	Theme	Information
Preceding – Anaphora	**3.796792**	**-3.58875**
Following	**-15.8481**	**14.97976**

Table 15 Contingency Table comparing Theme/Information clausal positions with type of following reference

	Theme	Information			χ^2 table
Within group – Esphora	*97*	*139*	236		**0.00257**
E =	**96.50202**	**139.498**			**0.055131**
Beyond group - Cataphora	*4*	*7*	11		**0.001778**
E =	**4.497976**	**6.502024**			**0.038139**
	101	146	247	χ2=	**0.097618**
				phi=	**0.01988**

Table 16 Contingency Table comparing Initial/Final clausal positions with Participant Tracking status

	Simple or Initial Complex Theme	Clause or Clause-Complex final			χ^2 table
Participant	574	592	1166		8.897316
E =	650.0507	515.9493			27.81306
Non-Participant	284	89	373		11.20983
E =	207.9493	165.0507			35.042
	858	681	1539	χ2=	82.96221
				phi=	0.232178

Table 17 Contingency Table comparing Initial/Final clausal positions with Referent/Addition

	Simple or Initial Complex Theme	Clause or Clause-Complex final			χ^2 table
No referent – Addition	*138*	*243*	381		**13.09513**
E =	**187.5592**	**193.4408**			**6.355726**
Referent	*436*	*349*	785		**12.69697**
E =	**386.4408**	**398.5592**			**6.162478**
	574	592	1166	χ2=	**38.3103**
				phi=	**0.181263**

Table 18 Contingency Table comparing Initial/Final clausal positions with Tracked Context

	Simple or Initial Complex Theme	Clause or Clause-Complex final			χ^2 table
Context of Culture – homophora	*9*	*32*	41		**8.328978**
E =	**22.77197**	**18.22803**			**0.458989**
Context of Situation	*427*	*317*	744		**10.40526**
E =	**413.228**	**330.772**			**0.573408**
	436	349	785	χ2=	**19.76663**
				phi=	**0.158683**

Table 19 Contingency Table comparing Initial/Final clausal positions with Lexical Relations

	Simple or Initial Complex Theme	Clause or Clause-Complex final			χ^2 table
Superordination	*349*	*251*	600	A1	**0.010396**
E =	**347.1004**	**252.8996**		A2	**0.014268**
Composition	*43*	*40*	83	B1	**0.52391**
E =	**48.01556**	**34.98444**		B2	**0.719058**
Nuclear Relations	*17*	*7*	24	C1	**0.699319**
E =	**13.88402**	**10.11598**		C2	**0.959803**
	409	298	707	χ^2 =	**2.926753**
				phi=	**0.06434**

Table 20 Contingency Table comparing Initial/Final clausal positions with Superordination type

	Simple or Initial Complex Theme	Clause or Clause-Complex final			χ^2 table
Co-Hyponymy	*324*	*238*	562		**0.025668**
E =	**326.8967**	**235.1033**			**0.379611**
Class-SubClass	*25*	*13*	38		**0.035689**
E =	**22.10333**	**15.89667**			**0.527826**
	349	251	600	χ2=	**0.968795**
				phi=	**0.040183**

Table 21 Contingency Table comparing Initial/Final clausal positions with Repetition/Substitution

	Simple or Initial Complex Theme	Clause or Clause-Complex final			χ^2 table
Complete Repetition	*158*	*153*	311		**3.579388**
E =	**183.6381**	**127.3619**			**10.21275**
Substitution	*90*	*19*	109		**5.160978**
E =	**64.3619**	**44.6381**			**14.72536**
	248	172	420	χ2=	**33.67847**
				phi=	**0.283173**

Table 22 Contingency Table comparing Initial/Final clausal positions with Preceding/Following reference

a. Contingency Table

	Simple or Initial Complex Theme	Clause or Clause-Complex final			χ^2 table
Preceding – anaphora	*338*	*283*	621		**1.256837**
E =	**359.2489**	**261.7511**			**9.075533**
Following	*71*	*15*	86		**1.724988**
E =	**49.75106**	**36.24894**			**12.45602**
	409	298	707	χ2=	**24.51338**
				phi=	**0.186205**

b. Percentage Deviation

	Simple or Initial Complex Theme	Clause or Clause-Complex final
Preceding – Anaphora	**-5.91482**	**8.117995**
Following	**42.71052**	**-58.6195**

Table 23 Contingency Table comparing Initial/Final clausal positions with type of following reference

	Simple or Initial Complex Theme	Clause or Clause-Complex final			χ^2 table
Within group – esphora	*70*	*11*	81		**0.146306**
E =	**66.87209**	**14.12791**			**2.37016**
Beyond group - cataphora	*1*	*4*	5		**0.692516**
E =	**4.127907**	**0.872093**			**11.21876**
	71	15	86	χ2=	**14.42774**
				phi=	**0.409591**

Appendix 4.1 FMAM Results for Participant, Theme & Information Systems

Key

underline = Theme
italics = *Information*
bold = **Presenting**
grey = Presuming
1 size larger text = Clause-Complex Internal Final Position
2 sizes larger text = Clause-Complex Final position

Modulation Techniques for Mobile Radio

Modulation is *the process of* ***encoding information from a message source in*** **a manner suitable for** ***transmission.*** It generally involves translating ***a baseband message signal*** (called *the source*) ***to a bandpass signal*** ***at*** **frequencies** **that** are *very high* when compared to *the baseband frequency.* The bandpass signal is called *the modulated signal* and the baseband message signal is called *the modulating signal.* **Modulation** may be *done* by varying *the amplitude, phase, or* frequency of ***a high frequency carrier*** *in accordance with the amplitude of the message signal.* **Demodulation** is *the process of* ***extracting*** *the baseband message from the carrier* so that it may be processed and interpreted by *the intended receiver* (also called *the sink*).

This chapter describes ***various modulation techniques*** that are used in ***mobile communication systems.*** **Analog modulation schemes** that are employed in **first generation mobile radio systems,** as well as **digital modulation schemes** proposed for **use in present and future systems,** are *covered*. Since **digital modulation** offers ***numerous benefits*** and is already being used to replace ***conventional analog systems,*** the primary emphasis of this chapter is *on* ***digital modulation schemes.*** However, since **analog systems** are ***in widespread use,*** and will continue to *exist*, they are treated *first.*

Modulation is ***a topic*** that is covered *in great detail* ***in various communications textbooks.*** Here, the coverage focuses on ***modulation and*** ***demodulation*** as it applies to ***mobile radio systems.*** **A large variety of modulation techniques** have been studied for ***use*** ***in*** **mobile radio communications systems.** **and** **research** is *ongoing.* Given the hostile fading and multipath conditions in the mobile radio channel, designing ***a modulation scheme*** that is resistant to ***mobile channel impairments*** is ***a challenging task.*** Since the ultimate goal of **a modulation technique** is to transport *the message signal through* ***a radio channel with*** *the best possible quality* while occupying *the least amount of* ***radio spectrum,*** **new advances in digital signal processing** continue to bring about ***new forms of*** *modulation and*

demodulation. This chapter describes ***many practical modulation schemes, receiver architectures, design tradeoffs. and*** *their performance under* **various types of *channel impairments.***

6.1 Frequency Modulation vs. Amplitude Modulation

Frequency modulation (FM) is *the most popular analog modulation technique* used in ***mobile radio systems.*** In FM, the amplitude of the modulated carrier signal is kept *constant* while its frequency is varied by *the modulating message signal.* Thus, FM signals have *all their information in the phase or frequency of the carrier.* As shown subsequently, this provides ***a nonlinear and very rapid improvement in reception quality*** once **a certain minimum received signal level,** called *the FM threshold* is *achieved.* In amplitude modulation (AM) schemes, there is ***a linear relationship between*** *the quality of the received signal and the power of the received signal* since **AM signals** superimpose *the exact relative amplitudes of the modulating signal onto the carrier.* Thus, **AM signals** have *all their information in the amplitude of the carrier.* FM offers ***many advantages over amplitude modulation*** *(AM)*, which makes *it* ***a better choice for many mobile radio applications.***

Frequency modulation has ***better noise immunity*** when compared to ***amplitude modulation***. Since **signals** are represented as ***frequency variations*** *rather than amplitude variations*, FM signals are *less susceptible to* ***atmospheric and impulse noise***. which tend to cause ***rapid fluctuations in*** *the amplitude of the received radio signal.* Also, **message amplitude variations** do not carry ***information in*** *FM,* so **burst noise** does not affect ***FM system performance as much as AM systems,*** provided that the FM received signal is *above the FM threshold.* **Chapter 5** illustrated how **small-scale fading** can cause ***rapid fluctuations in*** *the received signal,* thus FM offers ***superior qualitative performance in fading*** when compared to *AM*. Also, **in an FM system,** it is possible to tradeoff ***bandwidth occupancy for improved noise performance.*** Unlike AM, **in an FM system** the modulation index, and hence bandwidth occupancy, can be varied to obtain *greater signal-to-noise performance.* It can be shown that, **under certain conditions,** the FM signal-to-noise ratio improves *6 dB for each doubling of* ***bandwidth occupancy.*** This ability of **an FM system** to trade bandwidth for SNR is perhaps *the most important reason for its superiority over AM.* However, **AM signals** are able to occupy *less bandwidth as compared to FM signals,* since the transmission system is *linear.* In modern AM systems, **susceptibility to fading** has been dramatically improved through *the use of* ***in-band pilot tones*** which are transmitted along with *the standard AM signal.* The modern AM receiver is able to monitor *the pilot tone* and rapidly adjust *the receiver gain* to compensate for *the amplitude fluctuations.*

An FM signal is ***a constant envelope signal,*** due to *the fact that the envelope of the carrier does not change with* ***changes in*** *the modulating signal.* Hence the transmitted power of **an FM signal** is *constant regardless of the*

amplitude of the message signal. The constant envelope of the transmitted signal allows ***efficient Class C power amplifiers*** to be used for ***RF power amplification of*** *FM.* In AM, however, it is critical to maintain ***linearity between*** *the applied message and the amplitude of* *the transmitted signal,* thus **linear Class .A or AB amplifiers,** which are not as *power efficient,* must be *used.*

The issue of **amplifier efficiency** is *extremely important* when designing ***portable subscriber terminals*** since the battery life of the portable is tied to *the power amplifier efficiency.* **Typical efficiencies for Class C amplifiers** are *70%*, meaning that 70% of the applied DC power to the final amplifier circuit is converted into ***radiated RF power.*** **Class A or .AB amplifiers** have ***efficiencies on*** *the order of 30-40f%.* This implies that for the same battery, **constant envelope FM modulation** may provide *twice as much talk time as AM.*

Frequency modulation exhibits ***a so-called capture effect characteristic.*** The capture effect is ***a direct result of*** *the rapid nonlinear improvement in received quality for* ***an increase in received power.*** If **two signals in** the same frequency band are available **at** ***an FM receiver,*** the one appearing at the higher received signal level is *accepted and* *demodulated,* while the weaker one is *rejected.* This inherent ability to pick up *the strongest signal* and reject the rest makes ***FM systems*** *very resistant to* ***co-channel interference*** and provides ***excellent subjective received quality.*** In AM systems, on the other hand, all of the interferers are received *at once* and must be discriminated after *the demodulation process.* While FM systems have ***many advantages over*** *AM systems,* they also have ***certain disadvantages.*** FM systems require ***a wider frequency band in*** *the transmitting media* (*generally several times as large as* *that* *needed for* *AM*) in order to obtain *the advantages of* ***reduced noise and capture effect.*** **FM transmitter and receiver equipment** is also *more complex than* *that* *used by* *amplitude modulation systems.* Although frequency modulation systems are tolerant to ***certain types of signal and*** ***circuit nonlinearities,*** **special attention** must be given to ***phase characteristics.*** Both AM and FM may be *demodulated* using ***inexpensive noncoherent detectors.*** AM is easily *demodulated* using ***an envelope detector*** whereas FM is *demodulated* using ***a discriminator or slope detector.*** AM may be detected coherently with ***a product detector,*** and in such cases AM can outperform *FM in* ***weak signal conditions*** since FM must be received ***above threshold.***

Appendix 4.3 Sec Results for Participant, Theme & Information Systems

Key

underline	= Theme
italics	= *Information*
bold	= **Presenting**
grey	= Presuming
1 size larger text	= Clause-Complex Internal Final Position
2 sizes larger text	= Clause-Complex Final position

7.1 Introduction

Security measures must be incorporated into ***computer systems*** whenever they are ***potential targets for*** *malicious or mischievous attacks*. This is *especially so* ***for systems*** that handle ***financial transactions or confidential, classified or other information*** *whose secrecy and integrity* are *critical*. **In Figure 7.1,** we summarize *the evolution of* ***security needs in computer systems*** since they first arose with *the advent of* ***shared data in multi-user timesharing systems of*** *the 1960s and 70s*. Today the advent of **wide-area, open distributed systems** has resulted in ***a wide range of security issues.***

The need to protect the integrity and privacy of **information** and other resources belonging to **individuals and organizations** is pervasive *in both the physical and* *the digital world*. It arises from *the desire* ***to share resources.*** In the physical world, **organizations** adopt ***security policies*** that provide for *the sharing of resources within* ***specified limits.*** For example, **a company** may permit ***entry to*** *its buildings for its employees and for* ***accredited visitors.*** **A security policy for documents** may specify *groups of* ***employees*** who can access *classes of* ***documents*** or it may be defined for ***individual documents and users.***

Security policies are enforced with *the help of* ***security mechanisms.*** For example, **access to a building** may be controlled by ***a reception clerk,*** who issues ***badges to accredited visitors,*** and enforced by ***a security guard or by electronic door locks.*** **Access to paper**

documents is usually controlled by ***concealment and restricted distribution.***

In the electronic world, the distinction between **security policies and mechanisms** remains *important;* without it, it would be *difficult to determine* whether **a particular system** was *secure.* **Security policies** are *independent of the technology* used, just as the provision of a lock on a door does not ensure *the security of a building* unless there is ***a policy for*** *its use (for example, that the door will be locked* whenever nobody is guarding *the entrance).* The security mechanisms that we shall describe do not *in themselves* ensure *the security of a system.* **In Section 7.1.2,** we outline *the requirements for security* ***in various simple electronic commerce scenarios,*** illustrating *the need for* ***policies in*** *that context.* **As an initial example,** consider *the security of* ***a networked file server*** whose interface is accessible to ***clients.*** To ensure that access control to files is maintained, there would need to be ***a policy*** that **all requests** must include ***an authenticated user identity.***

The provision of mechanisms for the protection of data and other computer-based resources and for securing **networked transactions** is *the concern of this chapter.* We shall describe *the mechanisms* that enable ***security policies*** to be enforced in ***distributed systems.*** The mechanisms we shall describe are *strong enough* to resist *the most determined attacks.*

The distinction between security policies and security mechanisms is *helpful* when designing ***secure systems,*** but it is often *difficult* to be *confident* that **a given set of** security mechanisms fully implements *the desired security policies.* **In Section 2.3.3,** we introduced ***a security model*** that is designed to help in analysing *the potential security threats in* ***a distributed system.*** We can summarize *the security model of* ***Chapter 2*** *as follows:*

- Processes encapsulate *resources* (such as *programming language-level objects and other system-defined resources)* and allow ***clients*** to

access *them through their interfaces.* **Principals (users or** other processes) can be explicitly authorized to operate on ***resources.*** Resources must be protected against ***unauthorized access.***

- Processes interact through *a network* that is shared by *many users.*

Enemies (attackers) can access *the network.* They can copy or attempt to read ***any message*** transmitted through *the network* and they can inject ***arbitrary messages,*** addressed to ***any destination*** and purporting to come from ***any source,*** *into the network.*

That security model identifies *the features of distributed systems* that expose *them to attacks.* In this chapter, we shall detail *these attacks* and *the security techniques* that are available for defeating *them.*

Appendix 4.4 RM Results for Participant, Theme & Information Systems

Key

underline = Theme
italics = *Information*
bold = **Presenting**
grey = Presuming
1 size larger text = Clause-Complex Internal Final Position
2 sizes larger text = Clause-Complex Final position

RETARDATION METHODS
INTRODUCTION

1. **For many years** **fixed wing aircraft** relied purely on ***wheel brakes*** to retard *their forward movement* *after landing.* Over the years, **advancements in design** and the introduction of **new materials** has led to ***increased capabilities*** ***and carrying capacities,*** which has inevitably led to *faster and heavier aircraft.* Increased flying speeds have led in turn to *higher landing speeds,* which has increased *the burden* placed on *the wheel brakes.*

2. **Without some form of additional assistance** in retarding the aircraft, the wheel brakes would need to be applied in ***a heavy and sustained manner*** leading to ***rapid wear*** *of the brakes and tyres,* or runways would need to be extended ***to unacceptable lengths.***

To help reduce brake and tyre wear and to maintain runways **at an acceptable length, additional retardation methods** based on the principle **of aerodynamic braking** have been devised to supplement *the wheel brakes.*

3. Methods of **emergency retardation** have also been devised to arrest *the forward motion of the aircraft on the ground in the event* ***of failure*** *of the wheel brake system or* ***during an aborted take-off*** where **normal braking methods** would be *inadequate.*

WHEEL BRAKES

4. Wheel brakes form *the primary method of* retarding *the forward movement of* **most aircraft** *when on the ground* and, in common with **most braking systems,** they rely on *the principle of* energy conversion *for* their operation. In this method, **friction** is used to convert *the forward motion of* the aircraft ***(kinetic energy)*** *into heat,* which is then dissipated, *to the surrounding air.*

5. Wheel brakes and their associated operating systems are covered ***in Section 3 Chapters 1 and 3*** *of this Volume.*

AERODYNAMIC BRAKING

6. The term **aerodynamic braking** means using *the airflow* to assist in *retardation* by selectively increasing *the amount* ***of drag*** produced by *the aircraft.* This increase in drag is normally applied, in conjunction with *operation of the wheel brakes, immediately after landing or* ***during an aborted take-off*** and is produced by *one, or sometimes* ***a combination of*** *more than one, of the following methods:*

- **Flying control surfaces.**
- **Thrust reversal.**
- **Brake parachute.**
- **Angle of Attack.**

Flying control surfaces

7. The flying control surfaces which can be used to help retard *the aircraft* ***through aerodynamic braking*** *are:*

- **Airbrakes.**
- **Elevators**
- **Flaps.**
- **Tailerons**
- **Spoilers.**
- **Foreplanes**

8. As the airbrakes, flaps or spoilers are moved out *into the airflow* they cause ***an increase in*** *drag.* The further into the airflow they protrude, the greater the drag produced and, just as in flight, **any increase in** drag slows down *the forward progress of the aircraft.*

It should be noted that only the airbrake is specially designed to slow *the aircraft*; for the other controls this is ***only a secondary function.*** The flying control surfaces which provide ***aerodynamic braking*** are shown ***at Fig 1.***

Thrust reversal

9. Thrust reversal is used to assist in *aircraft retardation* by using ***engine power as a deceleration force.***

10. Propeller driven aircraft. On propeller driven aircraft, **thrust reversal** is achieved by reversing *the pitch of the propeller blades* so that the thrust generated by the propeller is directed *forwards instead of rearwards,* the degree **of braking assistance** being controlled by *use of the engine throttle lever(s).*

11. Turbo-jet aircraft. **Thrust reversal on turbo-jet aircraft** is achieved by changing

the direction of the exhaust gas stream through more than 90 degrees, but less than 180 degrees. **One method of deflecting** the exhaust gas stream is shown **at Fig 2.** During flight the bucket-type doors are held in the open position as shown **in Fig 2A,** allowing **rearward passage** of the exhaust gas stream. When braking assistance is required, the doors are moved into the gas stream **by hydraulic or pneumatic jacks** as shown **in Fig 2B,** uncovering apertures through which the gas is deflected **in a slightly forward direction** to give 'thrust reversal'.

Brake parachute

12. **Brake parachutes** are normally to be found on **fast jet aircraft** and produce **aerodynamic drag** to assist the wheel brakes in retarding the aircraft during its landing run.
A typical brake parachute installation consists of the parachute assembly **in housing with associated release/ jettison controls and mechanisms (Fig 3).**

13. The parachute assembly. The parachute assembly consists of the main parachute, the auxiliary (or drogue) parachute and the streamer cable. **A ring** on the free end of the streamer cable connects the parachute assembly to the airframe **through a release unit.** The drogue may be **a small, spring-loaded version** of the main parachute or connected to **a rigid conical cap,** which forms the tip of the a/c tail cone during flight. The main parachute **and streamer cable** are packed into **a canvas bag or light alloy canister (hopper),** whilst the drogue may be packed with the main parachute and streamer cable or may be located separately. The main parachute canopy is made of **vented nylon panels or ribbons,** connected to the streamer cable **by nylon rigging lines and a connector block**.

14.
The parachute housing. **A typical brake parachute housing** is located *as near as possible to* the end of the rear fuselage. The housing comprises of **a compartment** designed to accommodate the parachute assembly or hopper with **a spring loaded door** or **detachable cap,** which closes off the opening to the compartment.

15. The release and jettison controls and mechanisms. Control of the brake parachute is normally effected mechanically by **a Teleflex or Bowden control cable system,** but it may be achieved electrically or, **on some aircraft, by a combination of both methods. Cockpit selection** for both the release and

jettison of the brake parachute is usually made from ***a single control,*** *except for* ***combination systems*** where the parachute is released mechanically by ***a lever or handle*** **and** **a separate switch** is used to operate ***an electrical jettison unit.*** The cockpit control is normally guarded to prevent ***inadvertent operation.***

16. Principles of operation
When **'release' or 'stream'** is selected at *the cockpit control, the door or end cap of the compartment housing* the brake parachute opens to release *the drogue into the airstream* as shown in ***Fig 4A.***

17. The drag of the drogue parachute withdraws *the main parachute from its pack* ***along with a streamer cable,*** as the cable is pulled *taut* the main parachute develops (opens) to provide *the braking drag* as shown ***in Fig 4B.***

NOTES
(1) **On brake parachute installations fitted with a door,** the door is usually designed to hinge open *under the influence of* ***a damped spring*** when the parachute release control is operated, allowing *the spring loaded drogue parachute* to deploy and withdraw *the main parachute.*
(2) On those installations where the parachute housing is closed off by ***a cap,*** the cap is connected to *the drogue chute* ***by nylon cords.*** When the parachute release control is operated on *this type of installation,* the cap is completely released from *the airframe structure* and, as it is carried away by *the airflow,* it withdraws *the drogue chute* which withdraws *the main parachute.*

18. Once the aircraft has slowed to ***a speed*** where the brake parachute is *no longer effective,* it is jettisoned *onto the runway or taxiway* to be collected by ***a brake parachute recovery team.***

Angle of attack.
19. **On some aircraft types,** during landing, the pilot of the aircraft will hold *the nose of the aircraft* ***high*** *off the runway,* rolling along on *his main wheels* *only.* This high angle of attack produces ***a lot of aerodynamic drag,*** which assists *the retardation of the aircraft.*

EMERGENCY RETARDATION

20. **Large aircraft** normally have ***relatively low landing speeds*** and can be brought to ***a controlled standstill in an emergency*** either by ***aerodynamic braking alone*** *or,* ***on some airfields, by an area*** *at the end of the runway referred to* ***as a 'run off area'.*** This area is either ***a stretch of*** *rough ground designed to absorb the forward momentum of the aircraft,* ***or a shallow pit*** *filled with* ***pebbles,*** both types being designed to slow *the aircraft down* ***on entry*** without causing ***significant damage.***

21. **Small, fast jet aircraft** present ***more of a problem during emergencies*** *on landing due to their often high landing speeds and the use of shorter runways.* The undercarriage **of small aircraft** also tend to be less robust than *those* ***of large aircraft*** and their engine air intakes are normally *quite close to the ground,* two factors which make *the use of* ***'run off areas'*** *unsuitable due to the high risks* ***of seriously damaging*** *the undercarriage* ***and of ingesting debris into the engine air intake(s).*** To overcome *these problems,* **many small aircraft** are fitted with ***Arrestor hooks,*** which work in conjunction with *the Rotary Hydraulic Arrestor Gear (RHAG)* installed on *the runway.* **Many airfields** from which **fast jet aircraft** *operate* are also equipped with ***crash barriers*** *located at the ends of* the runway.

NOTE
As a last resort **some aircraft** utilise ***emergency undercarriage (u/c) up selection*** *to enable the underside of the a/c* to be used as ***a retard medium.***

The arrestor hook

22. The arrestor hook was adopted from *the Royal Navy,* who used *it as the primary method of retardation* when landing *aircraft* ***on aircraft carriers.*** The arrestor hook is ***a strong metal arm with*** **a hook** ***on*** **the end.** The arm is attached to ***a strong point*** *on the rear of the fuselage* and is lowered to engage with *the RHAG arrestor wire* stretched across *the runway.* **Control of an arrestor hook system** may be *mechanical, electrical or electro-hydraulic.*

23. Mechanically operated arrestor hooks. In these installations the arrestor hook is controlled from *the cockpit* ***by a handle or lever.*** The cockpit control is connected by ***a Teleflex or an open control cable system, to a release unit*** which normally holds *the hook in the retracted position.* When the cockpit control is pulled, the release unit is tripped and the hook is forced rapidly downwards, either by ***a spring jack (Fig 5A)*** **or** *by the inherent spring of the hook blade* ***(Fig 5B).*** The spring jack or **a torque tube** acting at the arm pivot stops *the hook* from bouncing as it

is trailed along *the ground* ***after release.***

24. Electrically operated arrestor hooks. The hook is held in *the up position by the jaws of the release unit.* When **a hook down selection** is made the solenoid of the release unit is energised which opens *the jaws* allowing *the hook* *to lower.* The jaws are locked around *the roller on the upper surface of* the hook *by the roller* operating *the jaws* when the hook is manually moved to *the up position.*

25. Indication. **All types of arrestor hook installation** usually have ***a cockpit indicator light,*** which illuminates when the arrestor hook is *released.* The light is energised by ***a micro-switch*** *on the release unit* and also serves as ***a warning light*** to indicate that the hook is not securely locked *up*. **On some aircraft** there is also ***a mechanical indicator*** *on the release unit* which normally lies flush with *the aircraft structure* but protrudes when the release unit is *unlocked*.

26. Raising the arrestor hook. On aircraft fitted with **mechanically or electrically operated hooks**, the hook usually has to be raised *manually by the ground crew* using ***a lifting rig*** ***or mechanical hoist.*** This is done after the cockpit control lever has been returned to *the 'hook up' position* to reset *the release unit* so that it is ready to receive and lock *the hook* *in the retracted position.*

NOTE
On installations where the arrestor hook has to be raised *manually,* it is *vital,* ***for safety reasons,*** to ensure that the hook is securely *'locked up'* before removing *the lifting force.* **Bulb filaments** can become *defective* therefore it would be wrong to rely on *the cockpit indicator light* *alone.* **A visual check** of the up-lock must always be made using *if necessary,* ***an extending inspection mirror*** ***and a torch.***

27. Safety devices. To prevent *the accidental release* *of the arrestor hook,* **a safety device** is fitted either on *the spring jack,* to stop *the jack* *from extending,* or **as a safety bar** fitted to *the aircraft structure* to hold *the hook* *in the up position.* This safety device will be fitted at ***all times*** whilst the aircraft is *on the ground,* only to be removed before *flight or* *for testing.* **A**

secondary safety pin may also be fitted to *the cockpit control* to prevent *it* from being *moved*. **A warning pennant** is attached to *each of the safety devices.*

Rotary Hydraulic Arrestor Gear (RHAG)

28. The RHAG is designed to arrest *aircraft* fitted with ***arrestor hooks*** *after* ***an aborted take off or an emergency landing with a wheel brake*** ***or*** **brake parachute failure.** There are ***two RHAG assemblies,*** normally positioned ***approximately 1300 feet*** (***400 metres)*** *from each end of the runway,* both of which may be engaged *from either landing direction.*

29. **A typical RHAG installation (Fig 6)** comprises of ***a cable*** suspended across *the runway a little above the surface,* ***and two energy-absorbing units, one*** located at *each side of the runway.* As shown, each end of the cable is connected, ***via a heavy nylon tape,*** *to one of the energy absorbing units* which is mounted on ***a concrete block.*** The tape is guided by ***a runway sheaf*** mounted on ***a second concrete block.*** Each energy-absorbing unit consists of ***a tape drum*** connected to ***a vaned rotor*** operating ***between two sets of stationary vanes*** ***in an oil filled chamber.***

30. Normally, the cable is over-run by ***an aircraft,*** but when its arrestor hook has been *lowered* and engages *the cable,* the pull on the tape causes *the tape drums and rotors* to *revolve.* The resulting oil turbulence created provides *the retarding force* by converting *the kinetic energy of the aircraft into heat.*

31. **An electrically powered retrieve system** is fitted close to *each energy absorption unit.* This is used to wind *the tape* ***back*** *onto the drum* and re-tension *the cable* after it has been disengaged from *the aircraft's arrestor hook.*

The crash barrier

35. The crash barrier (**Fig 7)** is designed to stop ***a small or medium sized turbo-jet aircraft*** which, because of **some emergency,** is about to overrun *the end of* the runway. The crash barrier is never used to stop ***a propeller driven aircraft.*** The barrier is sited at *the end of the runway* and takes *the form of* ***a fence*** made up of ***vertical nylon ropes*** attached to ***upper and lower***

steel cables.

36. The upper cable is spliced into *the lower cable at each end* *of the barrier* and the ends of the extended lower cable are taken, ***via shock absorber cables,*** ***to two rotary brake units.*** At each end of the barrier **a suspension cable** is attached to *the upper steel cable* ***by clamps*** ***and a shear pin.*** The suspension cables pass over ***hinged stanchions*** which are normally *lowered* so that the barrier lies *on the ground.* The stanchions are raised electrically ***by remote control*** to bring *the barrier* ***into a vertical position*** *as shown* when there is ***a possibility*** that it will be *required.*

37. When the barrier is *engaged,* its lower cable is overrun by *the aircraft's wheels* and its upper cable passes over *the cockpit canopy.* The nylon barrier ropes then fold over *the aircraft's wings,* pulling *the upper barrier cable* *downwards* causing *the pins* attaching *the suspension cables* to *shear.* The continuing pull on the barrier, transmitted via *the shock absorber cables,* pulls on ***brake cables*** wound on *the brake drums* *of the rotary brake units,* thus causing *the drums* to *rotate.* Rotation of the drums releases ***a compressed air supply*** to apply *the brakes* and effect ***a steady retardation*** ***of*** **the aircraft.**

Section 1 Chapter 1
RAF PTC CN 3787
1-1-6 06-528a/01/B50
06-528a/01/B50 1-1-7

Appendix 5 - General Resources

Below are links to appendix 5.

Appendix 5 contains the results of the text analysis across the systems of Participant Identification, Participant Tracking, Theme, Logical and Information status. These are contained within XML files on this CD-Rom. They can be accessed if you have the UAM Corpus Tools software installed on your computer.

The UAM Corpus Tool can be installed from this disc by following these instructions:

1. From this link right click and select 'Save as' or 'Save target as'. (Or follow this link and follow the instructions to download the latest version)
2. Save the file in a suitable place on your computer.
3. Open the file that you saved.
4. The file will install UAM Corpus Tools.
5. Start UAM Corpus Tools when it is installed.
6. Select 'Open Project'.
7. Locate "AllText02.ctpr" on this CD-Rom.
8. Open it using UAM Corpus Tools from "Open Project".

UAM Corpus Tools will then load all of the texts, schemes and analyses for this project. You can select any of them to investigate. If you want to make changes to any of the files you will need to copy them to your own computer. The files on the CD-Rom will not change.

UAM Corpus Tools was designed and is maintained by Mick O'Donnell. All details available on the website: http://www.wagsoft.com/CorpusTool The software is open-source and is freeware, and is distributed here with permission form the author on an as-is basis. No liability is assumed or implied.