Registerial cartography: context-based mapping of text types and their rhetorical-relational organization

Christian M.I.M. Matthiessen
ENGL, FH, The PolySystemic Research Group
Hong Kong Polytechnic University
christian.matthiessen@polyu.edu.hk

Abstract

This paper is concerned with one of the three types of variation inherent in language — viz. register variation, or variation in meaning according to context of use. It reports on a long-term research programme designed to map the registers that collectively make up a language using one parameter within the context of use as the starting point — the field of activity characteristic the context in which a text of a given register unfolds. I present a typology/ topology of fields of activity, and go on to show how different types of activity favour different logico-semantic relations in the global organization of texts instantiating different registers. I then also illustrate registerial variation in the lexicogrammatical realization of logico-semantic relations. The part of the long-term research I focus on here is thus concerned with registerial variation relating to the chain of realizations from context (field of activity) to semantics (logico-semantic relations), and from semantics (logico-semantic relations) to lexicogrammatical realizations (with particular attention to congruence, i.e. congruent vs. incongruent realizations). At the end of the paper, I suggest that registerial cartography is an integral part of the development of appliable linguistics, a synthesis approach to language transcending the thesis and antithesis pair of theoretical linguistics and applied linguistics.

Inherent variability of language

Language is inherently **variable**, languages are inherently variable: variability is part of the power of language — the power to adapt to socially very diverse and ever-changing contexts, at the same time contributing to the constant change. As languages evolve, they tend to remain stable because they are inherently variable, adapting to changing conditions of use; their stability is of a higher order: languages are **metastable**.

The inherent variability of languages poses a fundamental problem for any theories based on the assumption that languages are uniform and homogeneous; but it was recognized by Halliday and others in Systemic Functional Linguistics from the start of the development of the theory in the early 1960s; Halliday and others continued the Firthian tradition of conceiving of languages as **polysystemic**, as systems of systems. Firth (1935/ 1957: 29) had warned against conceiving of language in terms of unity:

> The multiplicity of social roles we have to play as members of a race, nation, class, school, club, as sons, brothers, lovers, fathers, workers, churchgoers, golfers, newspaper readers, public speakers, involves also a certain degree of linguistic specialization. Unity is the last concept that should be applied to language. Unity of language is the most fugitive of all unities whether it be historical, geographical, national, or personal. There is no such thing as *une langue une* and there has never been.

In early work, Halliday and his colleagues developed Firth's insight into language as a **system of variation** (e.g. Halliday, 1978: 156), in a sense providing a *synthesis* of the thesis of the unity of language and Firth's antithesis, his argument against this kind of unity (cf. Matthiessen, 1993: 222). According to this synthesis, languages are inherently variable, shading into one another just as dialects

do; and language is modelled as **a probabilistic system** (long before the advent of today's "probabilistic linguistics", as formulated in Bod, Hay & Jannedy, 2003, and used within "statistical natural language processing", Manning & Schütze, 1999). Thus variation can be — and has been — characterized in probabilistic terms within the overall theory of language as a probabilistic system (e.g. Halliday, 1959, 1978, 1991a,b, 1993; Nesbitt & Plum, 1988; Matthiessen, 1999, 2006, in press b).

Halliday and his colleagues originally recognized two broad kinds of variation — a familiar kind, **dialectal variation** (including sociolectal variation) and a less familiar but equally important one, **registerial variation**, drawing on Firth's notion of restricted languages (e.g. Halliday, McIntosh & Strevens, 1964; Gregory, 1967; Hasan, 1973; Ure & Ellis, 1977, and an early corpus-based investigation of Scientific English by Huddleston et al., 1968). These two varieties of language are glossed by Halliday (e.g. 1978: 35) as "variety according to the user" (dialect, or dialectal variety) and "variety according to use" (register, or diatypic variety); he writes (op cit.: 157):

> A dialect is any variety of a language that is defined by reference to the speaker: the dialect you speak is a function of who you are. In this respect, a dialect differs from the other dimension of variety in language, that of register: a register is a variety defined by reference to the social context — it is a function of what you are speaking. It seems to be typical of human cultures for a speaker to have more than one dialect, and for his dialect shifts, where they occur, to symbolize shifts in register. A 'standard' dialect is one that has achieved a distinctive status, in the form of a consensus which recognizes it as serving social functions which in some sense transcend the boundaries of dialect-speaking groups. This is often associated with writing — in many cultures the standard dialect is referred to as the 'literary [i.e. written] language' — and with formal education. Because of its special status, speakers generally find it hard to recognize that the standard dialect is at heart 'just a dialect' like any other.

To **dialect variation** and **register variation**, Halliday and his colleagues added a third kind of variation, **codal variation** ("semantic style"), based on Bernstein's notion of codes and linguistic corpus-based investigations (e.g. Hasan, 1973, 1989; Halliday, 1994). These three types of variation can be located according to two of the global dimensions of the organization of language, the **cline of instantiation** and the **hierarchy of stratification** (cf. Halliday, 1994) — represented diagrammatically here as Figure 1 (based on Matthiessen, 2007).
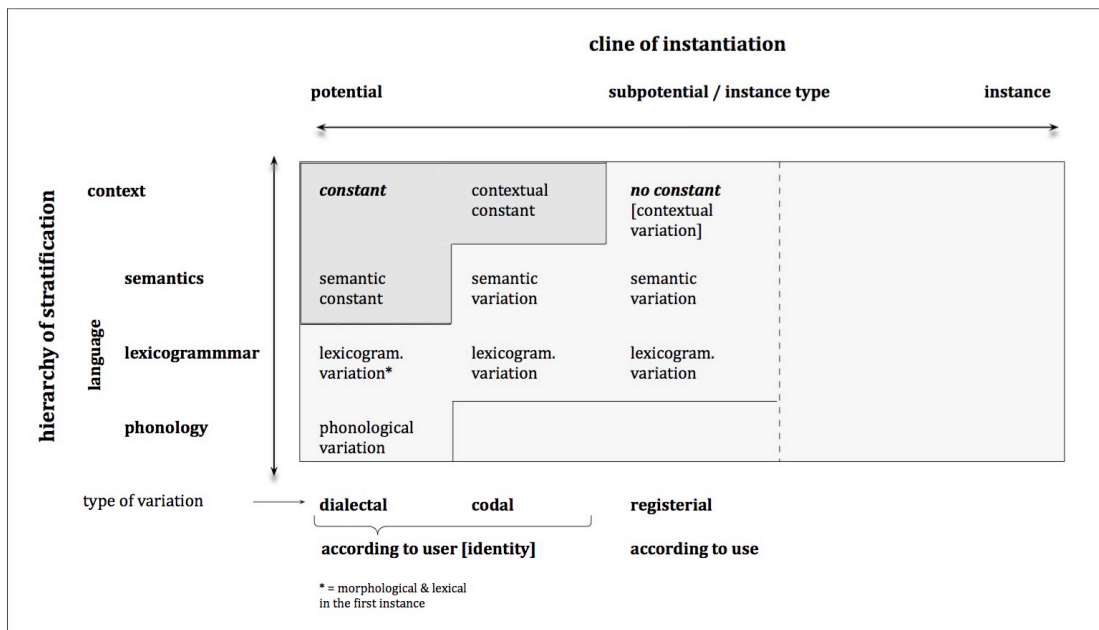

Figure 1: Locations of dialectal, codal and registerial variation along the cline of instantiation and the hierarchy of stratification — higher-level constant (if any) and primary nature of variation

The three types of variation are, in principle, distinct; but they interact in various ways, and have (as everything else in language) fuzzy boundaries — dialect variation obviously shading into language variation just as dialects shade into languages. As Halliday (1978) notes in the passage quoted above, different dialects may cover different **registerial ranges**, the standard dialect being an extreme example, as in the case of Standard English, which now embodies the registerial ranges collectively covered by English, Norman French and Latin before Standard English had evolved (cf. Halliday, 2003). Similarly, different codes are likely to embody different registerial ranges, reflecting both social hierarchy and the division of labour within a society.

Registerial variation

In this paper, among the three kinds of variation in Figure 1, I will be concerned with **registerial variation**. As shown in Figure 1, it is located **mid-region along the cline of instantiation**, between the potential pole of the overall (collective) meaning potential of a language and the instance pole of instantial acts of meaning unfolding to make up texts in context. In other words, we observe registerial variation (like any other kind of variation) as selections in texts as they unfold in their contexts of situation, and when we try to generalize these selections as recurrent patterns of selection, we find that the generalized patterns of selection are located mid-region along the cline of instantiation. In terms of stratification, it is **semantic variation** in the first instance, but it is semantic variation that co-varies with contextual variation: there is no higher-level constant, and this is precisely the notion of linguistic *variation according to use*, i.e. according to context of use. (In this important respect, registerial variation is unlike codal variation; codal variation is also semantic variation in the first instance [cf. Hasan, 1989, 2009], but it is variation with a contextual constant — codal varieties constitute different styles of meaning in comparable contexts, different semantic strategies for pursuing comparable contextual goals.) Registers are thus **meanings at risk**, describable as **probabilistic resettings** of the general systemic probabilities of a language (Halliday, 1978) operating within particular settings of contextual variables. They are distributed among the members of a speech community in terms of its division of labour; members — individual speakers — have different **registerial repertoires**, giving them access to different institutional roles.

Languages are aggregates of registers, and they evolve through registers. Registers emerge as adaptations to new contextual pressures on languages (as documented for the evolution of scientific English by Halliday, 1988, and as can be seen in the more recent evolution of e.g. news reporting and advertising, and now of course in the evolution of technologically enabled "electronic" registers), and they may fade away as contextual conditions change: the registerial make-ups of languages keep evolving, changing the character of languages in the course of evolution (cf. Halliday, 2013: Ch. 16).

Registers and register variation have been investigated, described and theorized since the 1960s — including the original Hallidayan version (in addition to the studies cited above, see e.g. Ure, 1982; Ghadessy, 1988, 1993; Teich, 1999; Steiner, 2004; and in computational modelling, e.g. Bateman & Paris, 1991) and US American register studies (e.g. Biber, 1988, 1995; Biber & Finnegan, 1994), with new insights coming from extensive text analysis and corpus-based studies; recent overviews include Lukin et al. (2008), Matthiessen (in press a) and also the introduction to the US American work on register by Biber & Conrad (2009)[1]. Biber & Conrad provide a helpful review of terms and concepts, and differentiate "genre", "style" and "register". Interpreted in terms of a Hallidayan systemic functional model, these three are arguably simply different manifestations of register variation — different in terms of the overall stratal and metafunctional organization of language in context, but not different in terms of the fundamental notion of functional variation in language — variation according to context of use[2].

---

[1] Registers have also been studied under different names, e.g. "text type", "genre"; and in machine translation, researchers have used the term "sublanguage" (e.g. Kittredge, 1987).

[2] Biber & Conrad (2009: Section 1.1) write of "the style perspective": "The key difference from the register perspective is that the use of these features is not functionally motivated by the situational context; rather, style features reflect aesthetic preferences,

Registerial cartography

Here I will report on aspects of a long-term project I have called **registerial cartography** (e.g. Matthiessen, in press a, forthc. b) — using the metaphor of cartography since those of us involved in the project are engaged in developing comprehensive *maps of registers* in different languages. These maps are based, in the first instance, on a "contextual projection": we approached registers "from above" (or "top down"), moving from context to semantics in terms of the hierarchy of stratification[3], adopting a view of them based on contextual parameters (variables), in particular on the three major parameters firs proposed by Halliday, McIntosh & Strevens (1964) and developed since then — field, tenor and mode (using the terms adopted by Halliday, 1978):

- **field** (type of activity): what's going on in context — the field of activity, and the field of experience accompanying or created by the activity (also known as "subject matter", "topic", "domain");
- **tenor** (role relationships): who are taking part — the tenor of the relationship among the interactants in terms of their roles and relations (including institutional roles, status roles, contact roles, sociometric roles);
- **mode** (symbolic organization): the role played by language, other semiotic systems and social systems in context — the complementary contributions made by them in context, including channel (graphic / phonic) and medium (spoken / written).

The contextual approach to the development of maps of functional variation, of register variation, is motivated by the very nature of this type of variation: *variation according to context of use*. However, at the same time, a central objective of the project of registerial cartography is to examine, describe and theorize registers according to Halliday's **trinocular vision** (e.g. Halliday, 1978: 130-131, 1996; Halliday & Matthiessen, 2013: 48-49), supplementing the view "from above" — from contexts, with the views "from below" — from lexicogrammar and phonology (or graphology), and "from roundabout" — from the level of semantics itself, the level at which the variation takes place in the first instance (in terms of the "meanings at risk" in different contexts). In other words, the project of registerial cartography includes centrally stratal coverage in the account of registers, from the contexts in which they operate to the linguistic strata where their semantic patterns are realized; stratal coverage thus includes a chain of inter-stratal realizations: context to semantics, semantics to lexicogrammar, and lexicogrammar to phonology or graphology (cf. Figure 6 in the Conclusion).

Of the different aspects of the registerial cartography project, I will focus in particular on the investigation of correlations between (i) fields of activity characterizing different types of context (situation types) and (ii) the choice of semantic strategies for organizing text within the register associated with a given type of context, with semantic strategy in the sense of logico-semantic relation (rhetorical relation, conjunctive relation, discourse relation).

Context: field of activity

In terms of **context**, I will present part of our typology of **fields of activity** (e.g. Matthiessen, in press a; Matthiessen & Kasyap, 2014; Matthiessen & Teruya, 2015), with types of activity differentiated in two to three steps in delicacy. The primary types are eight in number (derived from an unpublished manuscript by Jean Ure), each with subtypes as shown by means of a radial diagram in Figure 2:

associated with particular authors or historical periods." But "aesthetic preferences" are actually also functional, only in a different way, as was brought out by work by Mukařovský (1948) in the Prague School on the "esthetic function" of language. Cf. also Hasan (1985). For different uses of the terms "genre" and "register" in SFL, see e.g. Matthiessen (1993, in press a, forthc. b).
[3] In a sense corpus-based investigations such Biber (1988, 1995) have tended to move in "from below", using lexicogrammatical patterns that can be the basis of automated analysis in large volumes of text — though taking note of "situational factors" (e.g. Biber & Conrad, 2009). The two moves are *complementary* as strategies adopted to describe registers and registerial variation; and they need to be linked up through a chain of inter-stratal realizations (cf. Figure 9 below).

expounding (general classes of phenomena), **reporting** (particular instances of phenomena, typically chronicling events), **recreating** (some aspect of experience, imaginatively), **sharing** (personal values and experiences), **doing** (collaborating in, or directing, social behaviour), **enabling** (typically some course of action — some form of doing), **recommending** (some course of action or some commodity), **exploring** (assigning public value to commodities or arguing about ideas). These eight primary types of field of activity are characterized in Table 1, together with their immediate subtypes. Like all contextual and linguistics categories, fields of activity are indeterminate, and they shade into one another (see Matthiessen & Teruya, 2015).

Table 1: Primary and secondary fields of activity

primary type	nature of activity	secondary type
expounding	our experience of classes of phenomena according to a general theory (ranging from commonsense folk theories to uncommonsense scientific theories) —	either by **categorizing** (or "documenting") these phenomena (typically entities) or
		by **explaining** them (typically events or the outcomes of events);
reporting	on our experience of particular phenomena (instances of classes of phenomena), documenting them according to the principle of organization most salient to them (e.g. as a verbal time line, a verbal map or simply as a list) —	**chronicling** the flow of particular events (as in historical recounts or news reports),
		surveying particular places (as in guide books) or
		inventorying particular entities (as in catalogues);
recreating	our experience of the world imaginatively, that is, creating imaginary worlds having some direct or tenuous relation to the world of our daily lives — recreating the world imaginatively through	**narration** and/ or
		dramatization;
sharing	our personal lives, prototypically in private, thereby establishing, maintaining and negotiation personal relationships in face-to-face interaction but increasing also through social media channels (thus blurring the distinction between private and public) [sharing is a field of activity oriented towards tenor (relationships) so tenor distinctions play a significant role)] —	sharing our personal **experiences**, as in reminiscences, anecdotes and/ or
		sharing our personal **values**, as in gossip;
doing	social activities, prototypically engaging in interactive social behaviour, thereby collectively achieving some task —	either by members of one group **collaborating** with one another or
		by one person **directing** the other members of a group;
enabling	people to undertake some activity, thus very likely foreshadowing a 'doing' context —	either by **instructing** them in how to undertake the activity, as in 'how-to' manuals, or
		by **regulating** their behaviour (controlling, constraining and restricting it), as in legislation, contracts, licensing agreements;
recommending	people to undertake some activity, thus very likely	either by **advising** them

primary type	nature of activity	secondary type
	foreshadowing a 'doing' context —	(recommendation for the benefit of the addressee, as in professional consultations) or
		inducing them (promotion: recommendation for the benefit of the speaker, as in advertisements);
exploring	our communal values and positions, prototypically in public (through media channels) [exploring is a field of activity oriented towards tenor (relationships and values) so tenor distinctions play a significant role)] —	either by **reviewing** a commodity (goods-&-services), as in book reviews, or
		by **arguing** about positions and ideas, as in expositions, editorials, debates.
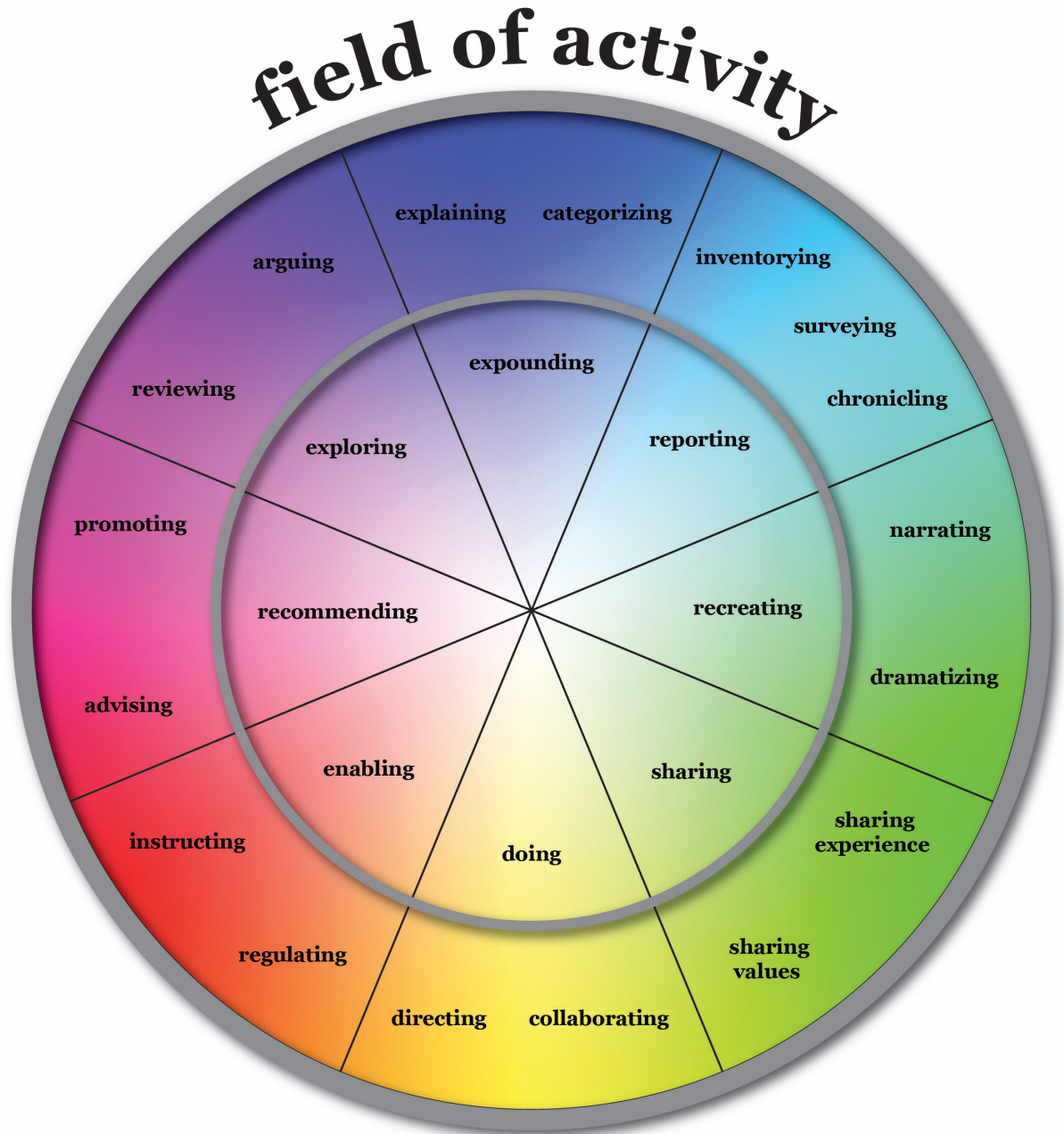
Figure 2: Context — the contextual parameter of field ("what's going on"): field of activity (the socio-semiotic process people are taking part in in context), primary types (inner circle) and secondary types (outer circle)

The description of field of activity diagrammed in Figure 2 and summarized in Table 1 includes two steps in delicacy — the eight primary types and their immediate subtypes; but it has of course been extended further in delicacy, and when we reach tertiary or quaternary delicacy in the differentiation of fields of activity, we can begin to relate the description to the categories of genre identified by systemic functional linguists working with Martin's (e.g. 1992) "genre model" — the genres of written language described by Martin & Veel (2008) and of spoken language described by Eggins & Slade (2005). These descriptions include the contextual structures of the genres, e.g. the structures of argumentative expositions and of explanations: see Table 2. The table contrasts sequential explanations with expositions (in the sense arguments supporting a thesis): we can specify the structure of both at the fourth step in delicacy in the

description of field of activity[4]. The two types are illustrated by Text 1, a sequential explanation, and Text 2, a(n analytic) exposition; for the sake of brevity, I have selected short educational texts of around ten clause complexes (orthographic sentences; for longer examples, see Matthiessen, forthc. a). The elements of their contextual structures are indicated in bold within square brackets; their logico-semantic structures will be presented below.

Table 2: Examples of differentiation of fields of activity in delicacy to the point where contextual structures can be posited

field of activity				contextual structure
primary	**secondary**	**tertiary**	**quaternary**	
expounding	explaining	sequentially	temporal	Phenomenon Identification ^ Explanation Sequence
exploring	arguing	one-sided	exposition	Thesis ^ Argument$_{1-n}$ ^ Reinforcement of Thesis

Text 1: Sequential explanation from an educational resource website[5] (structural conjunctions in bold, cohesive ones in bold italics)

[0] Woodchipping

[Phenomenon Identification:]

[1] Woodchipping is a process [[used to obtain pulp and paper product from the forest]]. [2] About 10 percent of Australia's state owned forest land, **and** large areas of privately owned forest, are involved in woodchip projects.

[Explanation Sequence:]

[3.1] The woodchipping process begins [3.2] **when** the trees are cut down in a selected area of the forest [[called a coupe]]. [4.1] *After that*, the tops and branches are cut off [4.2] **and** the logs are dragged to a log landing [4.3] <u>where</u> they are loaded onto a truck. [5.1] *Next* the bark of the logs is removed [5.2] **and** the logs are taken to a chipper [5.3] <u>which</u> cuts them into small pieces [[called woodchips]]. [6.1] The woodchips are *then* screened [6.2] to remove dirt and other impurities. [7.1] <u>At this stage </u> the woodchips are **either** exported to Japan in this form [7.2] **or** converted into pulp by chemicals, heat and pressure. [8.1] The pulp is *then* bleached [8.2] **and** the water content removed. [9.1] *Finally* it is rolled out [9.2] to make paper.

Text 2: Exposition ("analytical exposition") from an educational website[6]

[0] Cars should be banned in the city

[Thesis:]

[1] Cars should be banned in the city. [2.1] **As** we all know, [2.2] cars create pollution, [2.3] **and** cause a lot of road deaths and other accidents.

[Arguments:]

[3.1] *Firstly*, cars, << [3.2] **as** we all know,>> contribute to most of the pollution in the world. [4] Cars emit a deadly gas [[[that causes illnesses such as bronchitis, lung cancer, || **and** 'triggers' off asthma]]]. [5] Some of these illnesses are so bad [[that people can die from them]].

---

[4] The table only serves as a simple illustration. We may need to take further steps in delicacy, e.g. in order to distinguish analytical expositions (the type in focus here) from hortatory expositions, which include a recommendation for action to be taken based on the argument. In addition, we also need to take into consideration variations due to tenor, e.g. variation according to intended readers or listeners, and to mode, e.g. variation according to medium — spoken or written.
[5] http://www.schools.nsw.edu.au/media/downloads/schoolsweb/studentsupport/programs/lrngdificulties/writespellsec5.pdf
[6] http://sman5yk.sch.id/2013-03-21-17-03-23/inggris/232-english-lesson-material-for-grade-xi-semester-1

[6] *Secondly*, the city is very busy. [7.1] Pedestrians wander everywhere [7.2] **and** cars commonly hit pedestrians in the city, [7.3] which causes them to die. [8] Cars today are our roads' biggest killers.

[9] *Thirdly*, cars are very noisy. [10.1] **If** you live in the city, [10.2] you may find it hard to sleep at night, [10.3] or concentrate on your homework, [10.4] and especially talk to someone.

[Reinforcement of Thesis:]

[11] *In conclusion*, cars should be banned from the city for the reasons listed.

Semantics: logico-semantic (rhetorical) relations

In terms of the **semantic strategy** used to organize texts within their contexts, I will focus on **logico-semantic relations**, or "rhetorical relations"[7], modelling them by means of a version of Rhetorical Structure Theory (RST) — an approach to the semantic organization of text in terms of rhetorical relations that Bill Mann, Sandy Thompson and I started to develop a little over three decades ago, now sometimes referred to as "classical RST" (see e.g. Mann & Thompson, 1987; Matthiessen & Thompson, 1989; Mann & Matthiessen, 1991; Mann, Matthiessen & Thompson, 1992; Taboada & Mann, 2006; and for the use of RST in computational discourse processing, see e.g. Marcu, 1997, 2000; Carlson & Marcu, 2001 [RST annotation of documents from the Penn Treebank]; and cf. Stede, 2012,). The version I use here is a "systemicized" one, i.e. a version that differs from classical RST in that it is integrated within the overall SFL framework as a logical-semantic resource — with systemic organization as primary and structural organization as secondary, derived from the systemic organization by means of realization statements (see Matthiessen, forthc. a). The system is represented informally in Figure 3; this is a description of the resources in English for organizing texts relationally.

---

[7] Such relations have been investigated under many names including "conjunctive relations", "discourse relations", "rhetorical predicates", "coherence relations", "interpropositional relations".
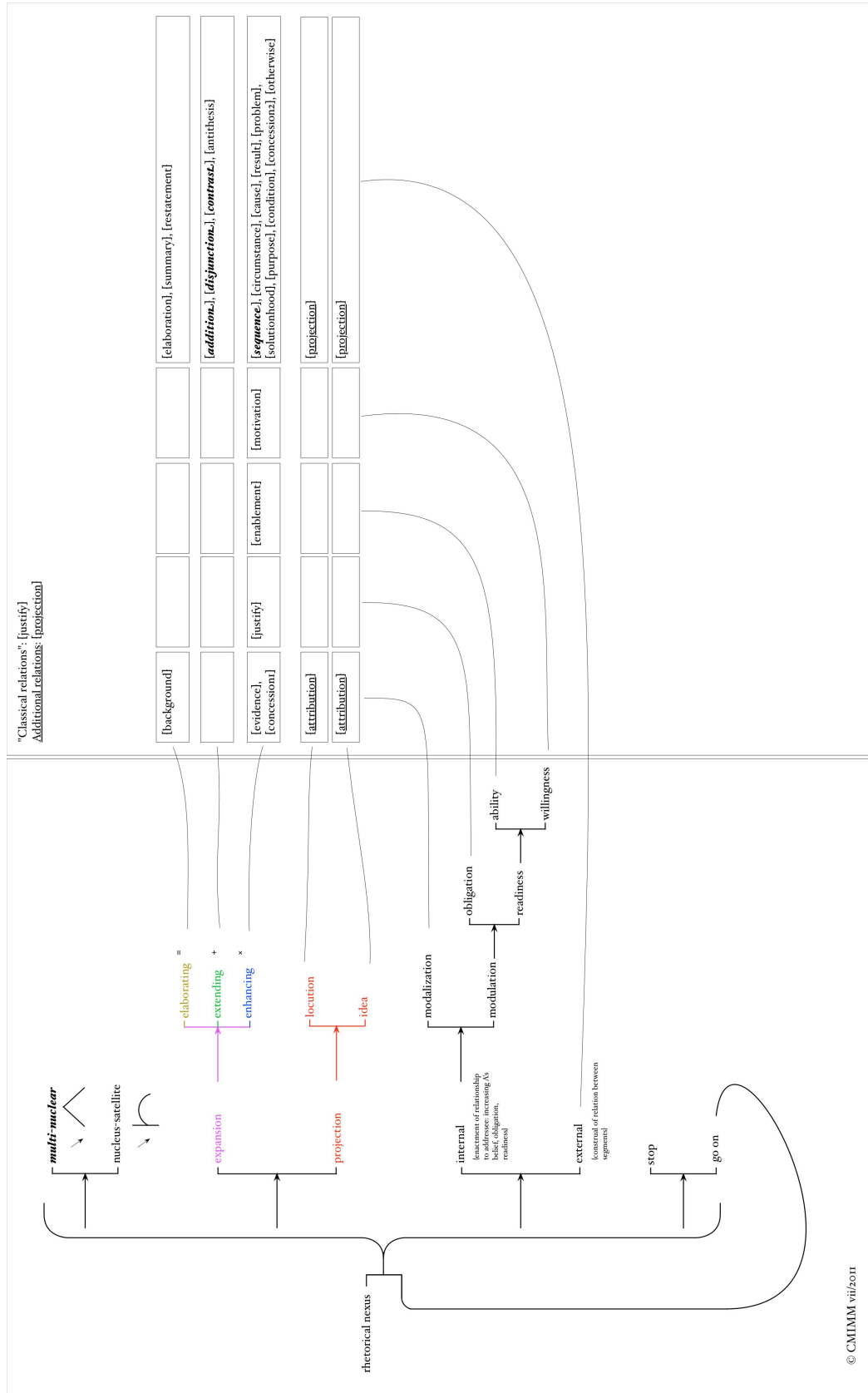
Figure 3: The semantic system of LOGICO-SEMANTIC RELATION (rhetorical relations)

The system of LOGICO-SEMANTIC RELATION in Figure 3 is composed of three simultaneous systems concerned with the nature of the logico-semantic relation used to relate one text segment to another in order to form a rhetorical nexus (i.e. a relational combination of text segments):

- The system of NUCLEARITY is the choice between relations linking the text segments as equal in status ('multi-nuclear') or as unequal, with one text segment supporting the other ('nucleus-satellite'). This distinction is part of "classical RST".
- The system of LOGICO-SEMANTIC TYPE is the choice between relations of 'projection', where one text segment sets up another as a quote or a report, and 'expansion', where one text segment elaborates, extends or enhances the other — the account of projection and expansion goes back to Halliday (1985).
- The system or orientation is the choice between linking two text segments as representations of experience ('external') or as interactional moves ('internal') — a distinction that goes back to Halliday & Hasan's (1976) description of cohesive conjunctions ("discourse markers") in English.

As can be seen from the table to the right of the system network in Figure 3, options (terms) from these three systems intersection to define sets of logico-semantic relations, including the "rhetorical relations" of classical RST. The relations can be fully differentiated if we increased the delicacy of the systems of LOGICO-SEMANTIC TYPE and ORIENTATION. For example, the relations marked by *finally* in Text 1 and *in conclusion* in Text 2 are similar in terms of LOGICO-SEMANTIC TYPE, both being enhancing relations, but different in terms of orientation: *finally* marks an 'external' relation whereas *in conclusion* marks an 'internal' one: see the logico-semantic analyses of these two texts in Figure 4 and Figure 5.

In addition to these three systems that jointly determine the nature of the relation linking the two text segments in a rhetorical nexus, there is a fourth system, the system of SYSTEMIC RECURSION. This is the choice between stopping the development of the text at the point of the current rhetorical nexus and going on to introduce a new logico-semantic relation thereby developing the text further.
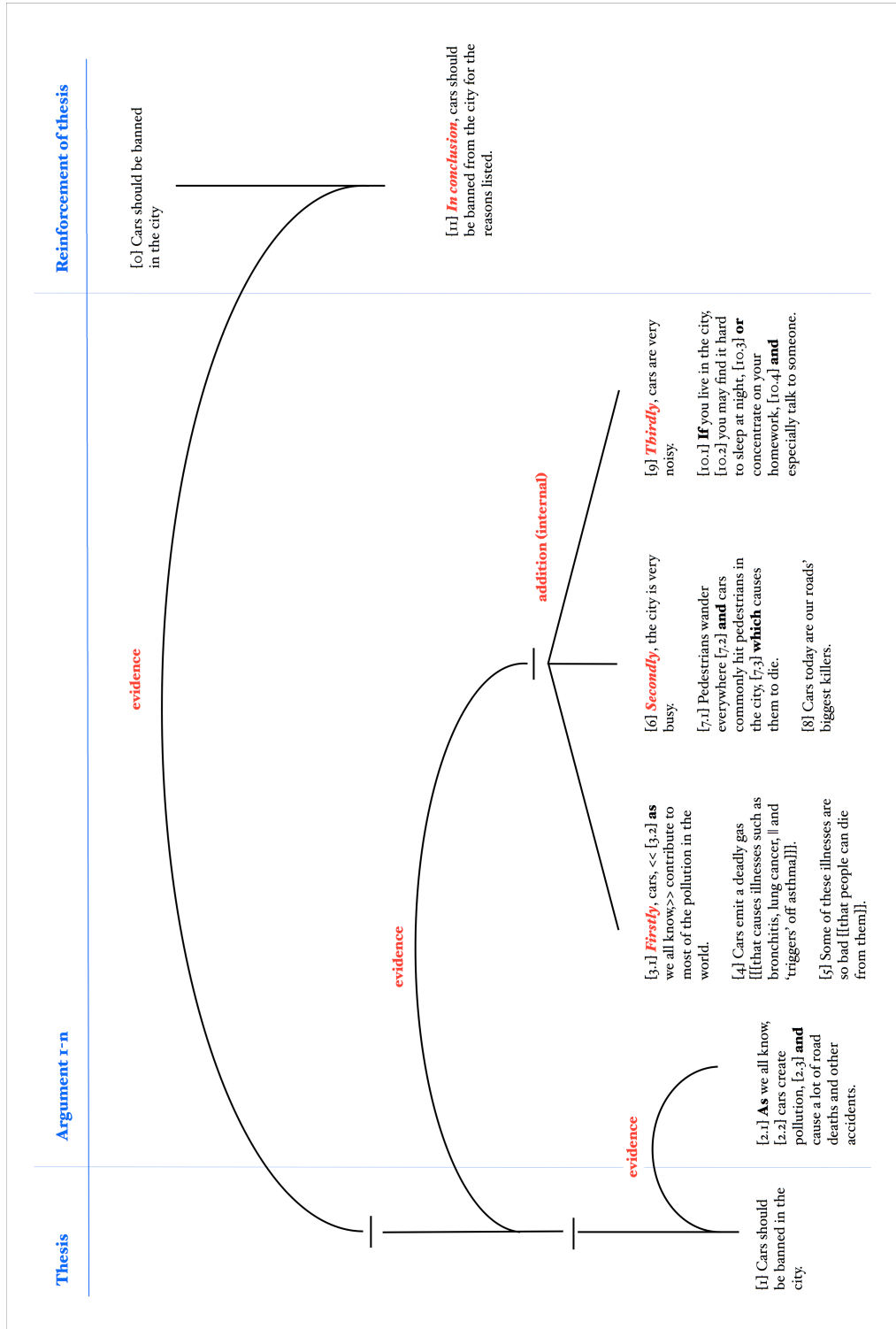
Figure 4: Logico-semantic analysis (in terms of RST) of the sequential explanation in Text 1

Figure 5: Logico-semantic analysis (in terms of RST) of the analytical exposition in Text 2

Fields of activity and favoured logico-semantic relations

Using the systemic description of logico-semantic relations in the organization of text set out in Figure 3, I have analysed representative samples of texts (mostly in English) from registers operating in
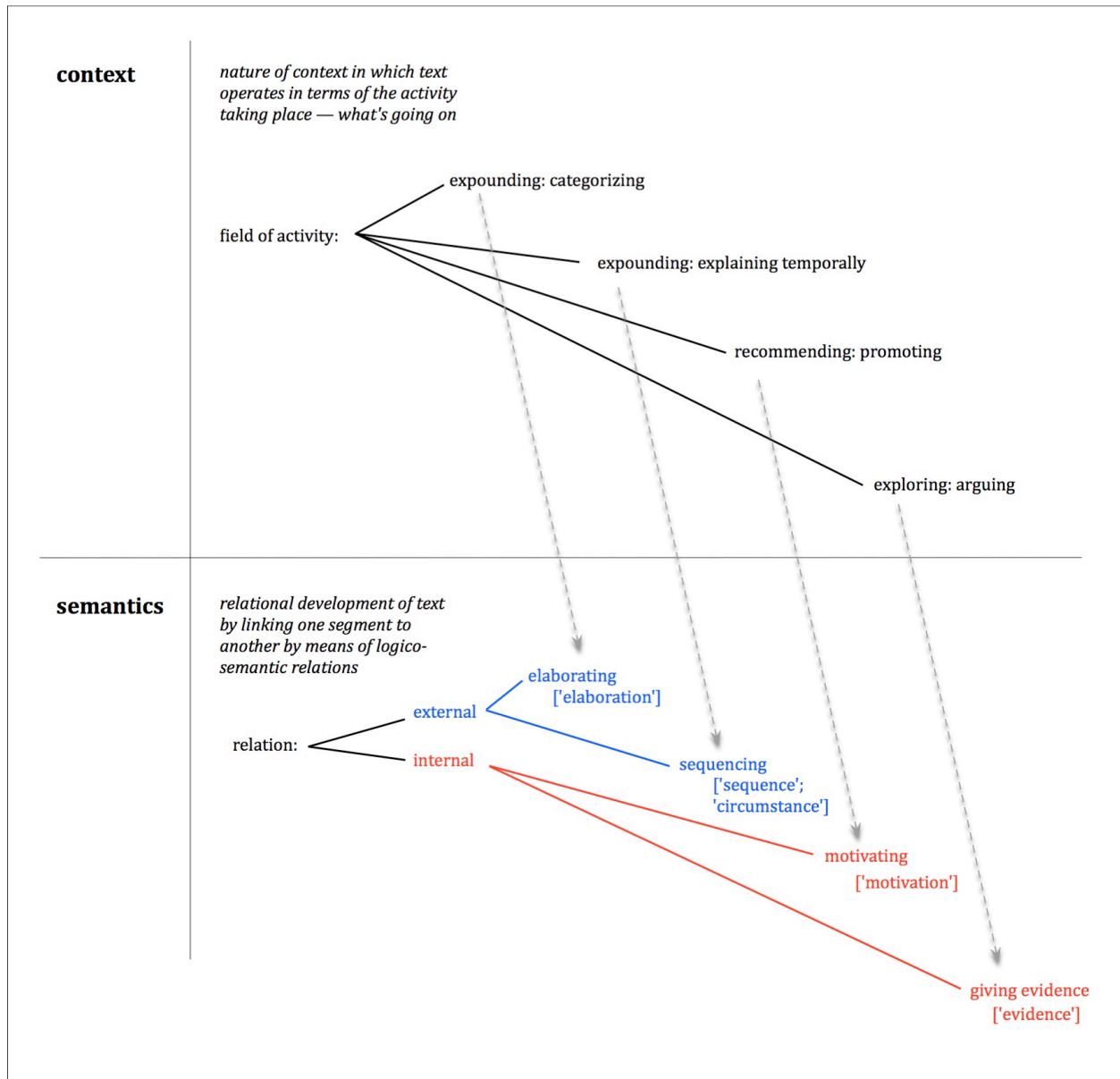
contexts characterized by different fields of activity. These analyses show, not surprisingly, that in the **global** organization of texts, different logico-semantic (rhetorical) relations are favoured (i.e. are "at risk" of being selected) according to the types of the field of activity characterizing the contexts in which the texts operate (see Matthiessen, in press a, forthc. a). This correlation between field of activity and logico-semantic relation becomes discernable when we increase the delicacy in the description of fields of activity from the eight primary types to their subtypes. As we differentiate these forms of activity further, identifying secondary and tertiary types (secondary types are shown above in the outer circle in Figure 2 and identified in the rightmost of column of Table 1), we can begin to discern recurrent semantic strategies used to organize texts belonging to registers operating in contexts characterized by one type of field of activity or other, as exemplified in Figure 6[8].

For example, if the field of activity of the context is one of expounding general knowledge by categorizing phenomena in terms of classes and subclasses or wholes and parts, the context will be realized by a taxonomic report where the key semantic strategy for organizing the text is the logico-semantic (rhetorical) relation of 'elaboration'; but if the activity is one of promoting some "commodity", the context will be realized by a marketing text such as an advertisement where the key semantic strategy for organizing the text is likely to be the logico-semantic relation of 'motivation', the point being to motivate the addressee to accept whatever is being offered.

Similarly, explaining phenomena by reference to the unfolding of processes in time will favour the logico-semantic relation of 'temporal sequence' as in Text 1, whereas arguing for a position or idea will favour the logico-semantic relation of 'evidence' as in Text 2. Thus the body of Text 1, which is an elaboration of the nuclear definition of 'woodchipping', is organized externally by means of multi-nuclear relations of 'sequence', as shown in Figure 4 above. In contrast, Text 2 is organized internally by means of nucleus-satellite relations of 'evidence', as shown in Figure 5 above. The satellite segments related by 'evidence' serve to bolster the writer's nuclear claim that cars should be banned in the city. The nucleus of the whole text comes at the end — as the culmination after the arguments in favour of the position it represents. This organization of expositions and other persuasive texts is typical — the global nucleus is presented as the "macro-New" of the whole text, the main point for readers or listeners to take away from the text.

---

[8] As noted above and illustrated in Table 2, this is roughly where contextual or situational structures — "generic structures", "schematic structures" — such as narrative structures begin to be identified and described: see Matthiessen (forthc. b) on the link to genre types identified and described by Martin & Rose (2008).

Figure 6: Examples of fields of activity (secondary types in Figure 2) with typical realizations by logico-semantic (rhetorical) relations playing role in organizing texts globally

The general principle is this: the meaning potential of a language, in this case of English, includes strategies for organizing texts by means of logico-semantic relations; and a certain subset of these will be most likely to be used (to be "at risk" of being chosen) in the global organization of texts in a context characterized by a particular type of field of activity. Different fields of activity will favour different subsets of relations. This general principle of registerial variation in the area of logico-semantic organization of text is represented diagrammatically in Figure 7. (Given a representative corpus texts from different registers that has been annotated for logico-semantic relations — cf. Carlson & Marcu, 2001, and Prasad et al., 2011, we would be able to state "favour" in probabilistic terms based on relative frequencies in the corpus.)
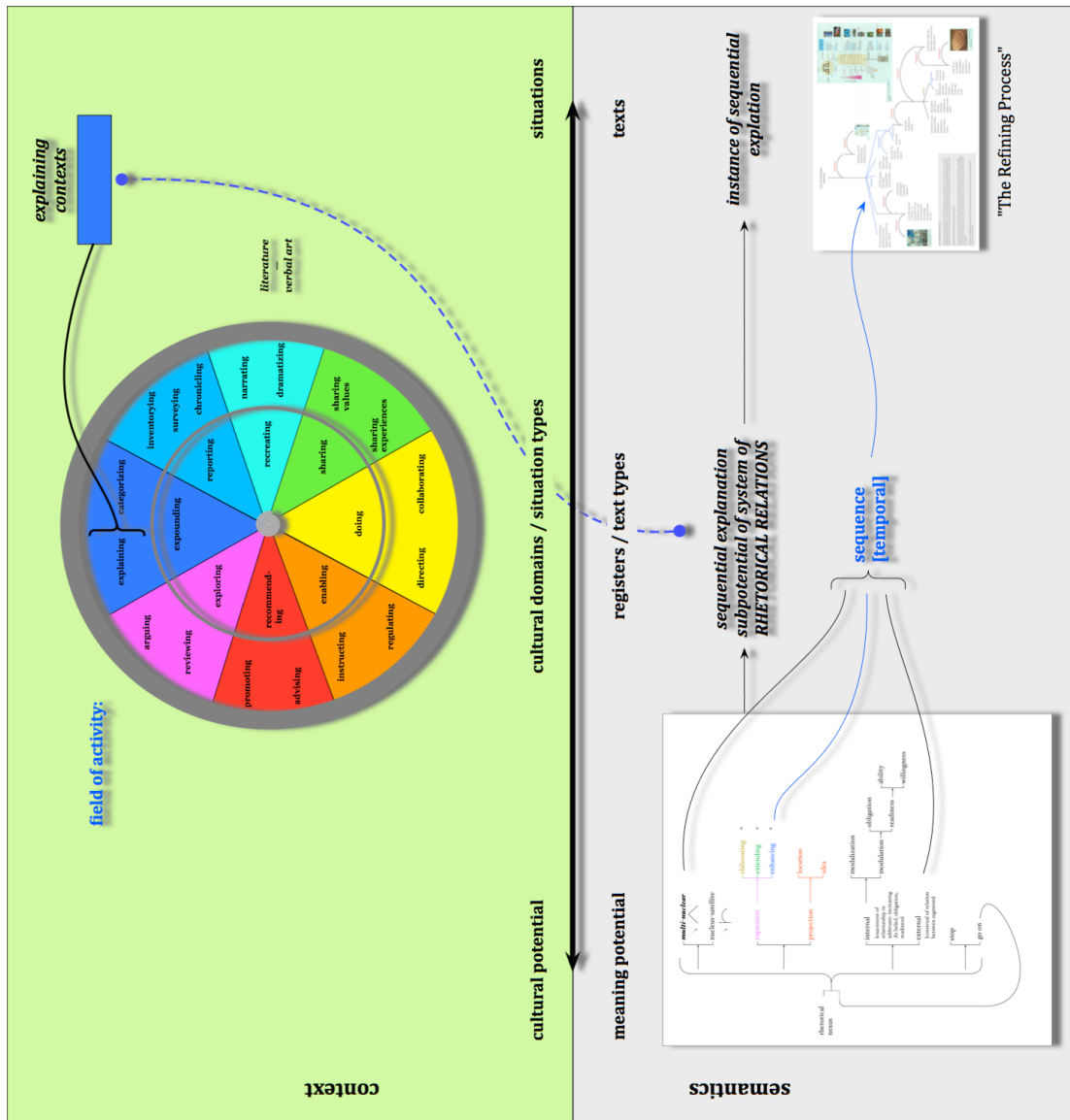
Figure 7: Registerial variation in the use of logico-semantic relations in the organization of texts belonging to different registers in accordance with the nature of the field of activity in context — exemplified by the activity of explaining by means of sequential explanations

Registerial variation in the lexicogrammatical realizations of logico-semantic relations

The logico-semantic relations favoured in the global organization of text thus vary according to the nature of the field of activity in context. By another step along the realizational chain from context to semantics and from semantics to lexicogrammar, we can also note that the **lexicogrammatical realizations** of logico-semantic relations similarly vary according to the nature of the field of activity (Matthiessen & Teruya, 2013). One interesting aspect of this variation in realization is the degree to which logico-semantic relations are realized **congruently** or **metaphorically** (incongruently). In texts of a **pragmatic** nature such as procedural texts operating in instructing enabling contexts (see the radial diagram in Figure 7), logico-semantic relations are likely to be realized congruently by conjunctions ("discourse markers"), either cohesive ones (e.g. *meanwhile*) or structural ones (e.g. *then*, *until*; *if*); but in texts of **mathetic** nature such as factorial explanations operating in expounding contexts, logico-semantic

relations are likely to be realized incongruently by prepositions (e.g. *because of*), verbs (e.g. *cause*, *lead to*, *result in*) or (by yet another step) nouns (e.g. *cause*, *consequence*, *effect*), as illustrated in Figure 8.
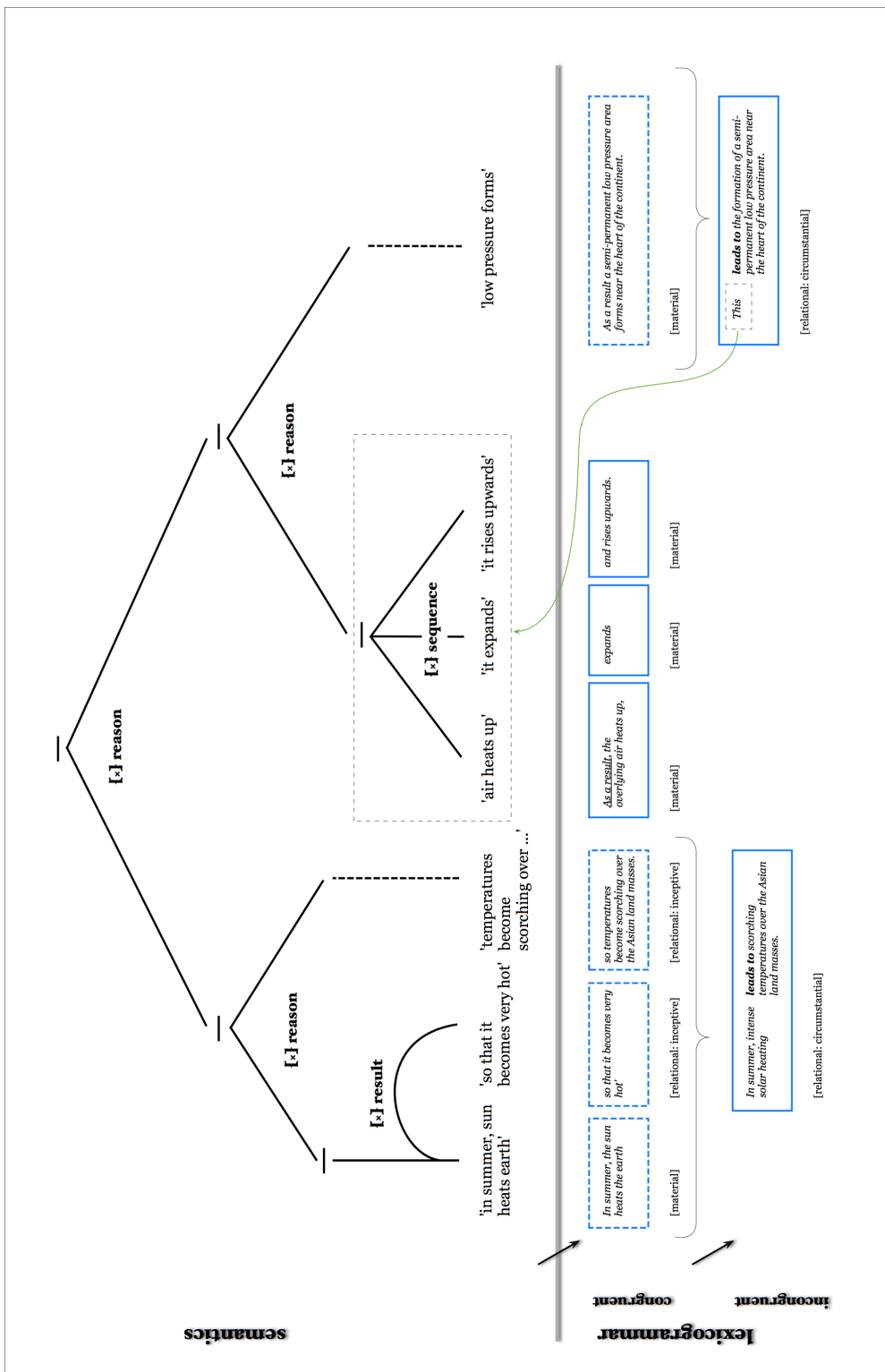

Figure 8: Congruent and incongruent realizations of logico-semantic relations in a passage from a causal explanation of monsoons

The text segment analysed in Figure 8 is an excerpt from a causal explanation of monsoons. It is organized by logico-semantic relations of 'reason', 'result' and (temporal) 'sequence', all of which are 'external' in orientation. The complex formed by relations of 'sequence' is realized congruently by a paratactic clause complex consisting of three 'material' clauses ("action" clauses). In contrast, the semantic complexes formed by means of 'reason' and 'result' are realized incongruently, by two 'circumstantial' 'relational' clauses, both of which have the causal verb *lead to* as Process. These incongruent clauses are as it were metaphoric re-codings of what would congruently be clause complexes, as indicated in Figure 8.

The metaphorical mode of realization has been investigated and discussed extensively in SFL based on Halliday (1985: Ch. 10), as in Halliday & Martin (1993), Halliday (1998), Vandenbergen, Taverniers & Ravelli (2003), Halliday & Matthiessen (2006: Ch. 6; 2013: Ch. 10), and modelled computationally as a feature of certain registers by Bateman & Paris (1991). Naturally, in addition to field of activity, other contextual parameters also play a role in shifting the realization of logico-semantic relations and rhetorical nexuses from the congruent mode to the metaphorical mode of realization; the metaphorical mode is more likely in written medium than in spoken, and, in terms of ontogenesis, more likely the further learners move into the subject-specific knowledge of secondary school (see e.g. Derewianka, 1995; Christie & Derewianka, 2008). Consequently, the realization of rhetorical nexuses is gradually "pushed down" in the lexicogrammar from cohesive sequences of clauses and clause complexes to clauses, phrases and groups. Incongruent, metaphorical realizations cover an important range of what Prasad, Joshi & Webber (2010) have identified as "alternative lexicalizations" ("AltLex") of "discourse relations" — alternative to (in our terms) congruent realizations in the form of structural and cohesive conjunctions.

Conclusion

In summary, I have reported on aspects of our research into registers — our long-term research programme of registerial cartography. In particular, I have discussed the relation between fields of activity within context, logico-semantic relations used to form rhetorical nexuses in the (global) semantic organization of text, and the mode of the lexicogrammatical realizations of these relations. This realizational chain is set out in Figure 9. The work discussed in this paper is exploratory, largely based on my manual analysis of samples of text that I have deemed to be representative of different registers. To scale up the research, one would need a registerial range of annotated corpora comparable to the discourse annotated version of the Penn Treebank (Prasad et al., 2007, 2008) and the more recent addition of the biomedical discourse relation bank (Prasad et al., 2011) — or one of the comparable corpora now becoming available for other languages, including Czech, Turkish and Hindi. With the aid of such a registerial range of corpora, or a single multi-registerial corpus, we would be able to check the patterns emerging in the exploratory work, scaling up the database to the point where statistically interesting statements can be made about the probabilistic settings of each register represented in the data — as a model, cf. Webber (2009) characterization of register varieties (in her terms, "genre distinctions") within the Penn Treebank.

The work on registerial cartography is, of course, important for its own sake: it sheds light on the essential nature of language as a system of variation — as an inherently variable, probabilistic system. In addition, there are many important areas of application where information registerial variation can lead to significant breakthroughs; these areas include education, translation, machine translation, computational discourse processing in general, multimodal studies (cf. Bateman, 2008; Matthiessen, 2009). In this way, registerial cartography is an integral part of **appliable linguistics** (cf. Halliday, 2008; Matthiessen, 2014a, 2014b).

context: field of activity   *expounding: explaining: sequential*

semantics: favoured rhetorical relations for   *expanding: enhancing: temporal: sequence*
                    forming global nexuses

lexicogrammar: favoured mode of realization   *[congruent] cohesive & tactic 'then'*
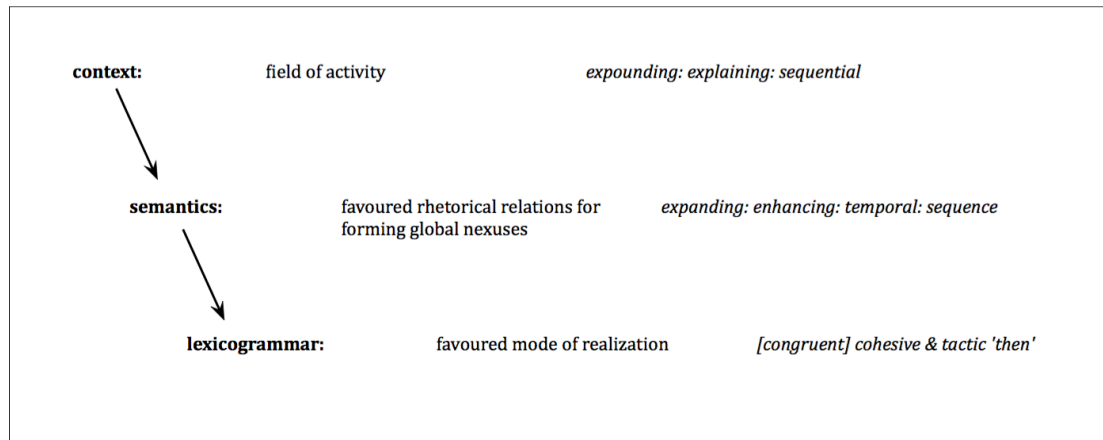
Figure 9: Stratification — the realizational chain discussed here

References

Bateman, John A. 2008. *Multimodality and genre: a foundation for the systematic analysis of multimodal documents.* London & New York: Palgrave Macmillan.

Bateman, John & Cécile Paris. 1991. "Constraining the deployment of lexicogrammatical resources during text generation: towards a computational instantiation of register theory." In Eija Ventola (ed.), *Functional and systemic linguistics: approaches and uses.* Berlin & New York: Mouton de Gruyter. 81-106.

Biber, Douglas. 1988. *Variation across speech and writing.* Cambridge: Cambridge University Press.

Biber, Douglas. 1995. *Dimensions of register variation: a cross-linguistic comparison.* Cambridge: Cambridge University Press.

Biber, Doug & Susan Conrad. 2009. *Register, genre, and style.* Cambridge: Cambridge University Press.

Biber, Douglas & Edward Finegan (eds.). 1994. *Sociolinguistic perspectives on register.* Oxford: Oxford University Press.

Bod, Rens, Jennifer Hay & Stefanie Jannedy (eds.). 2003. *Probabilistic linguistics.* Cambridge, Mass: MIT Press.

Carlson, Lynn & Daniel Marcu. 2001. *Discourse Tagging Reference Manual.* USC/ Information Sciences Institute.

Christie, Fran & Beverley Derewianka. 2008. *School discourse: Learning to write across the years of schooling.* London & New York: Continuum.

Derewianka, Beverly. 1995. *Language development in the transition from childhood to adolescence: the role of grammatical metaphor.* Macquarie University: Ph.D. thesis.

Eggins, Suzanne & Diana Slade. 2005. *Analysing casual conversation.* (First published by Cassell in 1997.) London: Equinox.

Firth, J.R. 1935. "The technique of semantics." *Transactions of the Philological Society.* Reprinted in In J.R. Firth (1957), *Papers in linguistics 1934-1951.* London: Oxford University Press. 7-33.

Ghadessy, Mohsen (ed.) 1988. *Registers of written English: situational factors and linguistic features.* London: Pinter.

Ghadessy, Mohsen (ed.). 1993. *Register analysis: theory and practice.* London & New York: Pinter.

Gregory, Michael J. 1967. "Aspects of varieties differentiation." *Journal of Linguistics* 3: 177-198.

Halliday, M.A.K. 1959. *The language of the Chinese "Secret History of the Mongols".* Oxford: Blackwell. (Publications of the Philological Society 17.) Reprinted in M.A.K. Halliday. 2006. *Studies in the Chinese language.* Volume 8 in the Collected Works of M.A.K. Halliday, edited by Jonathan J. Webster. London & New York: Continuum. 3-171.

Halliday, M.A.K. 1978. *Language as social semiotic: the social interpretation of language and meaning.* London: Edward Arnold.

Halliday, M.A.K. 1985. *An introduction to functional grammar.* London: Edward Arnold.

Halliday, M.A.K. 1988. "On the language of physical science." In Ghadessy (ed.), 162-178. Reprinted in Halliday, M.A.K. 2004. *The language of science.* Volume 5 in the Collected Works of M.A.K. Halliday. Edited by Jonathan J. Webster. London & New York: Continuum. 140-158.

Halliday, M.A.K. 1991a. "Corpus linguistics and probabilistic grammar." In Karin Aijmer & Bengt Altenberg (eds.), *English corpus linguistics: studies in honour of Jan Svartvik.* London: Longman. 30-43. Reprinted in Halliday (2005), Chapter 4: 63-75.

Halliday, M.A.K. 1991b. "Towards probabilistic interpretations." In Eija Ventola (ed.), *Trends in linguistics: functional and systemic linguistics: approaches and uses.* Berlin & New York: Mouton de Gruyter. Reprinted in Halliday (2005), Chapter 3: 42-62.

Halliday, M.A.K. 1993. "Quantitative studies and probabilities in grammar." In Michael Hoey (ed.), *Data, description, discourse: papers on the English language in honour of John McH. Sinclair.* London: Harper Collins. 1-25. Reprinted in Halliday (2005), Chapter 7: 130-156.

Halliday, M.A.K. 1994. "Language and the theory of codes." In Alan Sadovnik (ed.), *Knowledge and pedagogy: the sociology of Basil Bernstein.* Norwood, N.J.: Ablex. 124-142. Reprinted in MA.K. Halliday (2007), *Language and society.* Volume 10 in the Collected Works of M.A.K. Halliday, edited by Jonathan J. Webster. London & New York: Continuum. Chapter 8: 231-246.

Halliday, M.A.K. 1996. "On grammar and grammatics." In Ruqaiya Hasan, Carmel Cloran & David Butt (eds.), *Functional descriptions: theory into practice.* Amsterdam: Benjamins. 1-38. Reprinted in Halliday, M.A.K. 2002. *On grammar.* Volume 1 of Collected Works of M.A.K. Halliday. Edited by Jonathan Webster. London & New York: Continuum. Chapter 15: 384-417.

Halliday, M.A.K. 1998. "Things and relations: regrammaticizing experience as technical knowledge." In J.R. Martin & Robert Veel (eds.), *Reading science: critical and functional perspectives on discourses of science.* London: Routledge. 185-235.

Halliday, M.A.K. 2003. "Written language, standard language, global language." *World Englishes* 22(4): 405-418. Also in Braj B. Kachru, Yamuna Kachru & Cecil L. Nelson (eds.), 2006, *The handbook of World Englishes.* Oxford: Blackwell. 349-365.

Halliday, M.A.K. 2005. *Computational and quantitative studies.* Volume 6 in the Collected Works of M.A.K. Halliday, edited by Jonathan Webster. London & New York: Continuum.

Halliday, M.A.K. 2008. "Working with meaning: towards an appliable linguistics." In Jonathan J. Webster (ed.), *Meaning in context: implementing intelligent applications of language studies.* London & New York: Continuum. 7-23.

Halliday, M.A.K. 2013. *Halliday in the 21st century.* Volume 11 in the Collected Works of M.A.K. Halliday, edited by Jonathan J. Webster. London: Bloomsbury Academic.

Halliday, M.A.K. & Ruqaiya Hasan. 1976. *Cohesion in English.* London: Longman.

Halliday, M.A.K., Angus McIntosh & Peter Strevens. 1964. *The linguistic sciences and language teaching.* London: Longman.

Halliday, M.A.K. & J.R. Martin. 1993. *Writing science: literacy and discursive power.* London: Falmer.

Halliday, M.A.K. & Christian M.I.M. Matthiessen. 2006. *Construing experience through meaning: a language-based approach to cognition.* (First published by Cassell in 1999.) London & New York: Continuum.

Halliday, M.A.K. & Christian M.I.M Matthiessen. 2013. *Halliday's introduction to functional grammar.* Fourth Edition. London: Routledge.

Hasan, Ruqaiya. 1973. "Code, register and social dialect." In Basil Bernstein (ed.), *Class, Codes and Control: applied studies towards a sociology of language.* Volume 2. London: Routledge & Kegan Paul. 253-292.

Hasan, Ruqaiya. 1985. *Linguistics, language and verbal art.* Geelong, Vic.: Deakin University Press.

Hasan, Ruqaiya. 1989. "Semantic variation and sociolinguistics." *Australian Journal of Linguistics* 9: 221-275.

Hasan, Ruqaiya. 2009. *Semantic Variation: Meaning in Society and Sociolinguistics.* Volume Two in the Collected Works of Ruqaiya Hasan. Edited by Jonathan Webster. London: Equinox.

Huddleston, Rodney D. , Richard A. Hudson, Eugene Winter & A. Henrici. 1968. *Sentence and clause in Scientific English: final report of O.S.T.I.* Programme. University College London: Communication Research Centre.

Kittredge, Richard. 1987. "The significance of sublanguage for automatic translation." In Sergei Nirenburg (ed.), *Machine translation: theoretical and methodological issues.* Cambridge: Cambridge University Press. 59-67.

Lukin, Annabelle, Alison Moore, Maria Herke, Rebekah Wegener, Wu Canzhong. 2008. "Halliday's model of register revisited and explored." *Linguistics and the Human Sciences* 4(2): 187-243.

Mann, William C. & Christian Matthiessen. 1991. "Functions of language in two frameworks." *Word* 42(3): 231-49.

Mann, William C., Christian M.I.M. Matthiessen & Sandra A. Thompson. 1992. "Rhetorical Structure Theory and Text Analysis." USC/ISI Report. Also in William C. Mann & Sandra A. Thompson (eds.), *Discourse Description: Diverse Linguistic Analyses of a Fund Raising Text.* Amsterdam: Benjamins. 39-78.

Mann, W. & Sandra A. Thompson. 1987. *Rhetorical Structure Theory: A Framework for the Analysis of Texts.* ISI/RS-87-185.

Manning, Christopher D. & Hinrich Schütze. 1999. *Foundations of statistical natural language processing.* Cambridge, Mass.: The MIT Press.

Marcu, Daniel. 1997. *The rhetorical parsing, summarization, and generation of natural language texts.* University of Toronto: PhD thesis.

Marcu, Daniel. 2000. "The rhetorical parsing of unrestricted texts: a surface-based approach." *Computational Linguistics* 26(3): 395–448.

Martin, J.R. 1992. *English text: system and structure.* Amsterdam: Benjamins.

Martin, J.R. & David Rose. 2008. *Genre relations: mapping culture.* London & Oakville: Equinox.

Matthiessen, Christian M.I.M. 1993. "Register in the round: diversity in a unified theory of register analysis." In Mohsen Ghadessy (ed.), *Register analysis: theory and practice.* London: Pinter. 221-292.

Matthiessen, Christian M.I.M. 1999. "The system of TRANSITIVITY: an exploratory study of text-based profiles." *Functions of Language* 6(1): 1-51.

Matthiessen, Christian M.I.M. 2006. "Frequency profiles of some basic grammatical systems: an interim report." In Susan Hunston & Geoff Thompson (eds.), System and corpus: exploring connections. London: Equinox. 103-142.

Matthiessen, Christian M.I.M. 2007. "The "architecture" of language according to systemic functional theory: developments since the 1970s." In Ruqaiya Hasan, Christian M.I.M. Matthiessen & Jonathan Webster (eds.), *Continuing discourse on language.* Volume 2. London: Equinox. 505-561.

Matthiessen, Christian M.I.M. 2009. "Multisemiotic and context-based register typology: registerial variation in the complementarity of semiotic systems." Eija Ventola & Arsenio Jesús Moya Guijarro (eds.), *The world shown and the world told.* Basingstoke: Palgrave Macmillan. 11-38.

Matthiessen, Christian M.I.M. 2014a. "Appliable discourse analysis." Fang Yan & Jonathan J. Webster (eds.), *Developing Systemic Functional Linguistics: theory and application.* London: Equinox. 135-205.

Matthiessen, Christian M.I.M. 2014b. "Appliable linguistics: the potential of registerial cartography." MS of plenary given at the 3rd Forum on Applied Linguistics, the Guangdong Foreign Studies University, Guangzhou, P.R.C., 7 December 2014.

Matthiessen, Christian M.I.M. in press a. "Register in the round: registerial cartography. *Functional Linguistics* 1(2).

Matthiessen, Christian M.I.M. in press b. "Halliday's probabilistic theory of language." In Jonathan J. Webster (ed.), The Continuum Companion to M.A.K. Halliday. London & New York: Continuum.

Matthiessen, Christian M.I.M. forthc a. *Rhetorical System and Structure Theory: the semantic system of logico-semantic relations.* Book MS.

Matthiessen, Christian M.I.M. forthc b. "Modelling context and register: the long-term project of registerial cartography." Manuscript of book chapter submitted to Leila Barbara & Sara Cabral (eds.) *Teoria Sistêmico-Funcional para brasileiros* (Systemic Functional Theory for Brazilians). PPGL: Programa de Pós-Graduação em Letras. Universidade Federal de Santa Maria - UFSM: Santa Maria, Brazil.

Matthiessen, Christian M.I.M. & Abhishek Kumar Kasyap. 2014. "The construal of space in different registers: an exploratory study." *Language Sciences* 45: 1–27.

Matthiessen, Christian M.I.M. & Kazuhiro Teruya. 2013. "Grammatical realization of rhetorical relations in different registers." Paper manuscript.

Matthiessen, Christian & Kazuhiro Teruya. 2015. "Registerial hybridity: indeterminacy among fields of activity." In Donna Miller & Paul Bayley (eds.), *Permeable contexts and hybrid discourses.* London: Equinox.

Matthiessen, Christian M.I.M. & Sandra A. Thompson. 1989. "The Structure of Discourse and "Subordination"." In John Haiman & Sandra A. Thompson (eds.), *Clause Combining in Grammar and Discourse.* Amsterdam: Benjamins.

Mukařovský, Jan. 1948. "The esthetics of language." Extract from *Kapitoly z české poetiky*, translated by Paul Garvin (1964), *A Prague School reader on esthetics, literary structure and style.* Washington, D.C.: Georgetown University Press. 31-69.

Nesbitt, Christopher N. & Guenter Plum &. 1988. "Probabilities in a systemic grammar: the clause complex in English." In Robin P. Fawcett & David Young (eds.), *New developments in systemic linguistics, vol. 2: theory and application.* London: Frances Pinter. 6-39.

Prasad, Rashmi, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi & Bonnie Webber. 2008. "The Penn Discourse TreeBank 2.0." In *Proceedings, 6th International Conference on Language Resources and Evaluation,* Marrakech, Morocco.

Prasad, Rashmi, Aravind Joshi & Bonnie Webber. 2010. "Realization of Discourse Relations by Other Means: Alternative Lexicalizations." *Coling 2010: Poster Volume*, 1023–1031.

Prasad, Rashmi, Susan McRoy, Nadya Frid, Aravind Joshi & Hong Yu. 2011. "The biomedical discourse relation bank." *BMC Bioinformatics* 12(188): 1-18.

Prasad, Rashmi, Eleni Miltsakaki, Nikhil Dinesh, Alan Lee, Aravind Joshi, Livio Robaldo & Bonnie Webber. 2007. *The Penn Discourse Treebank 2.0 Annotation Manual.* The PDTB Research Group.

Stede, Manfred. 2012. *Discourse processing.* San Rafael, CA: Morgan & Claypool Publishers.

Steiner, Erich. 2004. *Translated texts: properties, variants, evaluations.* Frankfurt.Main: Peter Lang.

Taboada, Maite & William Mann. 2006. "Rhetorical Structure Theory: Looking back and moving ahead." *Discourse Studies* 8(4): 423–459.

Teich, Elke. 1999. "System-oriented and text-oriented comparative linguistic research: cross-linguistic variation in translation." *Languages in Contrast* 2(2): 187-210.

Ure, Jean. 1982. "Introduction: approaches to the study of register range." *International Journal of the Sociology of Language* 35: 5-23.

Ure, Jean N. & Jeffrey Ellis. 1977. "Register in descriptive linguistics and linguistic sociology." In Oscar Uribe-Villegas (ed.), *Issues in Sociolinguistics.* The Hague: Mouton.

Vandenbergen, Anne-Marie, Miriam Taverniers & Louise Ravelli (eds.) 2003. *Grammatical metaphor: views from systemic functional linguistics.* Amsterdam: John Benjamins.

Webber, Bonnie. 2009. "Genre distinctions for Discourse in the Penn TreeBank." *Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP.* 674–682.