Chapter 1

Introduction

1.1 On AI, Computational Linguistics, ~~universe and everything~~

In 1950 ~~Allan~~ Turing in a seminal paper (Turing 1950) published in Mind was asking if "machines can do what we (as thinking entities) can do?" He questioned what intelligence was and whether it could be manifested in machine actions indistinguishable from human actions.

He proposed the famous "Imitation Game" also known as the "Turing test" in which a machine would have to exhibit intelligent behaviour equivalent or indistinguishable from that of a human. The test was stating the following rules. The machine (player A) and a human (player B) are engaged in a written *natural language* conversation with a human judge (player C) ~~which~~ has to decide whether each conversation partner is human or a machine. The goal of players A and B is to convince the judge (player C) that they are human.

This game underpins the question whether "a computer, communicating over a teleprinter, (can) fool a person into believing it is human?", moreover whether it can ~~generate~~ human(-like) cognitive capacities (Stevan Harnad 1992). Essential parts of such cognitive capacities and intelligent behaviour that the machine needs to exhibit are of course the linguistic competences of comprehension (or "understanding") and generation of "appropriate" responses (for a given input from the judge C).

Artificial Intelligence (AI) field was born from dwelling on Turing's questions. The term was coined by McCarthy for the first time in 1955 referring to the "science and engineering of making intelligent machines" (McCarthy et al. 2006).

The general ~~tendency~~ is to program machines to do with language what humans do. Various fields of research contribute to this goal. Linguistics, amongst others, contributes with theoretical frameworks systematizing and accounting language in terms of ~~for~~ morphology, phonology, syntax, semantics, discourse or grammar in general. In computer science ~~are developed~~ increasingly more efficient algorithms and machine learning techniques. Computational linguistics provides ~~ingenious~~ methods of encoding linguistically motivated tasks in terms of formal data structures and computation goal. In addition, specific algorithms and heuristics operating within reasonable amounts of time with satisfiable levels of accuracy are tailored to accomplish those linguistically motivated tasks.

Computational Linguistics (CL) mentioned in 1950 in the context of automatic translation~~s~~ (Hutchins 1999) of Russian text into English ~~stared~~ developing before the field of Artificial Intelligence. Only a few years later CL became a sub-domain of AI as an interdisciplinary field dedicated to developing algorithms and computer software for intelligent processing of text (leaving the very hard questions of intelligence and human cognition ~~somehow~~ aside ~~that up to now still need massive inputs on human mind from cognitive, psycho-linguistic and other related sciences~~). Besides *machine translation* CL incorporates a broader range of tasks such as *speech synthesis and recognition, text tagging, syntactic and semantic parsing, text generation, document summarisation, information extraction*, etc.

This thesis contributes to the field of CL and more specifically it is an advancement in *Natural Language Parsing* (NLP), one of the central CL tasks informally defined as the process of transforming a sentence into (rich) machine readable syntactic and semantic structure(s). Developing a program to automatically analyse ~~the~~ text in terms of such structures by involving computer science and artificial intelligence techniques is a task pursued for several decades and still continues to be a major challenge today. This is especially so when the target is ~~a~~ *broad language coverage* (?) and even more when the desired analysis goes beyond simple syntactic structures towards richer functional and/or semantic descriptions useful in the latter stages of *Natural Language Understanding* (NLU).

In computational linguistics, broad coverage natural language components now exist for several levels of linguistic abstraction, ranging from tagging and stemming, through syntactic analyses to semantic specifications. In general, the higher the degree of abstraction, the less accurate the coverage becomes and the richer the linguistic description the slower the parsing process is performed.

~~These~~ working components are already widely used to enable humans to explore and exploit large quantities of textual data for purposes that vary from the most theoretical ~~ones~~ such as understanding how language works or the relation between form and meaning, to very pragmatic purposes such as developing systems with natural language interfaces, machine translation, document summarising, information extraction and question answering systems ~~and that is just~~ to name a few.

These software programs originally were designed by and for ~~the~~ domain experts but over time the fruits of the technological advancement became available to millions of ordinary people. In a world as ours, where technology is ~~so~~ ubiquitous and pervasive in~~to~~ almost all aspects of our life, natural language processing and understanding becomes of great value and importance regardless whether it materializes as a spell-checker, (not so) clever machine translation, voice controlled car or phone and so on.

1.2 The goal of the thesis

This thesis aims at a reliable modular method for parsing unrestricted English text into a feature rich constituency structure using Systemic Functional Grammars (SFGs).

Before describing the parsing method, the following aspects need to be clarified ~~first:~~ the theoretical framework and its descriptive power, the depth and meaningfulness of the analysis, the computational complexity of the process and the level of accuracy and how it is measured. I address each of these aspects in the following chapters and before advancing any further I would like to illustrate through an example ~~of~~ what parsing ~~is.~~

1.3 An example of Systemic Functional analysis

Traditional linguistic teaches us how to ~~do the~~ syntactic analysis of a sentence. So let's consider Example 1 in order to perform one. ~~We~~ need to focus on clustering words together into constituents guided by the intuitive rule "which word stands closest to another one" within the sentence. The word "He" is closest to "gave" and it seems that they go together especially ~~that~~ if we try to group "He" and "the" ~~they do~~ sense at all. Then, using sequential proximity criteria, "gave" must be related to "the" but they do not stick together at all it is rather "the" and "cake" that are a unity. So "the" has a stronger relation to "cake" than "gave". But actually "cake" seems connected to "gave" and not the other way around, so the direction of connection seems to matter just as much as its strength. So we see that the sequential proximity sometimes indicated relatedness between words but often it does not. The strength and direction as well are crucial and tt is some sort of meaning making that governs the grouping process. This grouping into syntactic constituents can be expressed using bracketed notation as in Example 2.

(1) He gave the cake away.

(2) ((He) (gave) ((the) (cake)) (away) (.))

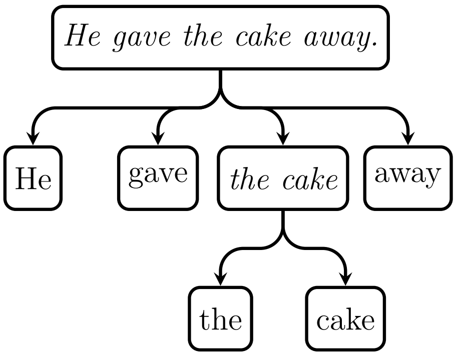


Fig. 1.1 Representation of the Example 2 as constituency tree

Figure 1.1 depicts the constituency division of the clause which is identical to the bracket notation in Example 2. The nodes represent grammatical constituents and the edges stand for the part-whole composition. This constitutes the simplest structure that could be expected from a parser without any specification of constituent class, function or any other grammatical functions. We know as well from works on parsing with formal generative grammars that such composition can always be expressed as a tree (or parse tree).

Each constituent can be decorated with its grammatical features. For example the word "he", we know, is third person pronoun, masculine gender and singular number; or the word "gave" is a verb and the predicate of the sentence, and so on. The structure in Figure 1.2 depicts a syntactic constituency tree in which every node is richly decorated with syntactic and semantic features. The blue part of each node represents grammatical class and function fundamental for establishing a valid constituency structure; the red part represents the semantic functions (called Transitivity in SFL); and the green part other grammatical features. In practice, the feature set is much richer than what nodes in Figure 1.2 carry, the current limitation being simply the space constraints. Generating automatically graphs like the one in Figure 1.2 is the purpose of the current work.

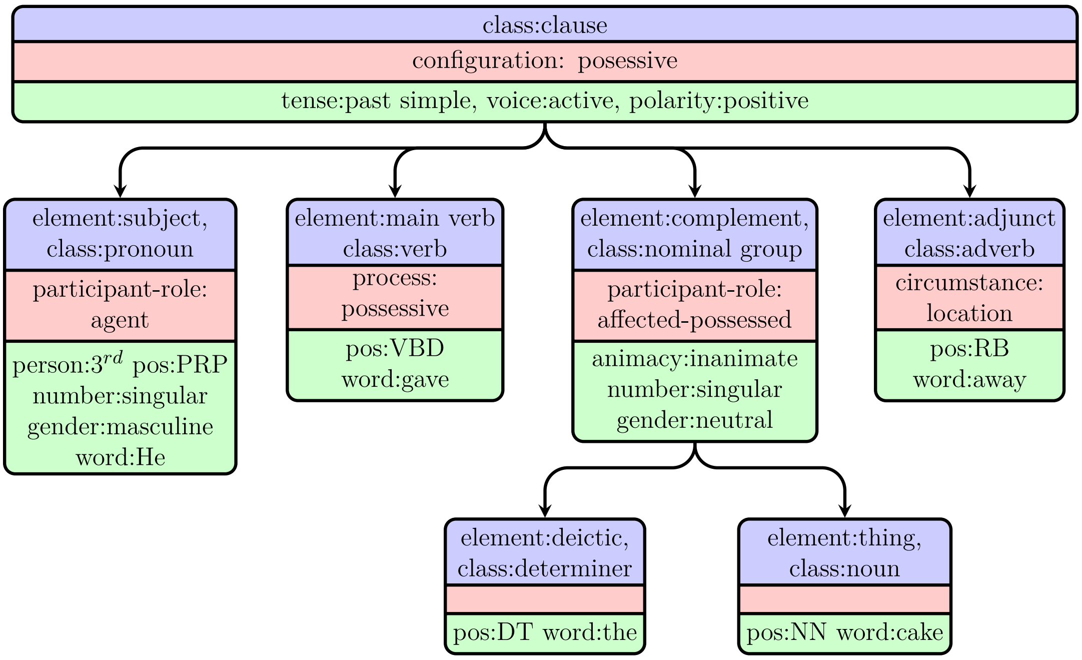


Fig. 1.2 Representation of Example 2 as feature rich constituency graph

1.4 The linguistic framework

In this thesis I chose a Systemic Functional Linguistic (SFL) stance because of its versatility in ~~describing and explaining in~~ *multiple semiotic dimensions* (Halliday 2003b) (i.e. paradigmatic, syntagmatic, meta-functional, stratification and instantiation dimensions) and at different *delicacy levels* of the *lexico-grammatical cline* (Halliday 2002; Hasan 2014) the (occurring) linguistic phenomena.

SFL regards language as a social semiotic system where any act of communication is regarded as a conflation of *linguistic choices* available in a particular language. Choices, ~~which~~ are organised on a paradigmatic rather than structural axis and ~~materialised~~ as *system networks*. Moreover, in SFL perspective language has evolved to serve particular *functions* influencing their the structure and organisation of the language. However, ~~being~~ around the paradigmatic ~~systems,~~ the functional organization differs significantly from other functional approaches (Prague School, Lexical Functional Grammar, Head-driven phrase structure grammar, Role and Reference Grammar, Functional Discourse Grammar and other). Halliday refers to the language functions as metafunctions or lines of meaning offering a trinocular perspective on language through *ideational*, *interpersonal* and *textual* metafunctions.

In SFL, language is first of all an interactive action serving to enact social relations under the umbrella of *interpersonal metafunction*.Then it is a medium to express the embodied human experience of ~~the~~ inner and outer worlds via *ideational metafunction*. Finally the two weave together into a coherent discourse flow whose mechanisms are ~~explained~~ through the *textual metafunction*.

To account for the complexity and phenomenological diversity of human language ~~the~~ SFL theory provides descriptions along *syntagmatic, (meta)functional, paradigmatic, stratification and instantiation axes.*

There are two models of SFG: *Sydney Grammar* (SG) Halliday & Matthiessen (2013) developed by Halliday and Matthiessen, the founding fathers of *Systemic Functional Linguistics* (SFL), and *Cardiff Grammar* (CG) Fawcett (2008), an extension and a simplification of Sydney Grammar. Each of the two grammars has advantages and shortcomings which I present in analyse and select based on theoretical soundness and suitability to the goals of the current project.

Cardiff and Sydney grammars had been used as language models in natural language generation projects within the broader contexts of social interaction. Some researchers attempted to reuse the grammars for the purpose of syntactic parsing within the borders of NL generation coverage, such as that of Kasper (1988), O'Donoghue (1991), O'Donnell (1993), Souter (1996), Day (2007).

1.5 The SFG complexity problem

Bateman (2008) thoroughly explains the reasons for such tremendous complexity after the attempts of Kasper (1988), Kay (1985), O'Donoghue (1991), O'Donnell (1993) and Day (2007), just a few ~~to mention~~, none of which managed to parse broad coverage English with full SFG ~~and~~ without aid of some sort. Each had to accept limitations either in grammar or language size and eventually ~~using~~ simpler syntactic trees as a starting point of the parsing process. So what is it about?

Automatic analysis of text can be seen as a problem of searching though the space of possible solutions for an appropriate or even optimal solution. Here we speak of the Systemic Functional Grammar as a linguistic resource that shapes the search space and the way it access to that space is available. The systemic lexicogrammar is organised paradigmatically and was ~~proven~~ to be a good structure for natural language generation task but it turns out to be unusable for the reverse problem, that of natural language analysis. The principal issues is that of handling the *search space* leading back to Halliday's question "How big is a grammar?" (Halliday 1966).

The size of the search space defined by a grammar depends on the number of system networks and on the kind of connectivity and cross-classification it provides.

For example given 50 system networks, the size of the search space lies somewhere between 51 and 2^{50}. This nevertheless is not such a big deal in the case of generation, as Halliday (1996) says that the "number of choice points [...] is actually rather small" as only few of the actual possibilities produced by a system network need to be explored when generating a clause. "Possible feature selections become relevant only when they are revealed to be relevant by prior paradigmatic choices and it is only those alternatives that need to be considered"(Bateman 2008).

Analysis is not symmetric with generation ~~and~~ the paradigmatic context of choice that is available during generation is no longer accessible in parsing. It is not known any longer which features of a systemic network are relevant and which are not. That is: in generation, the simple traversal of the network finds only the compatible choices because that is what the network leads to; whereas in analysis it is not evident in advance which path to follow therefore the task is virtually to explore entire search space in order to discover which features apply to the text (Bateman 2008).

In the analysis task first difficulty that needs to be addressed is discovering from a sequence of words what possible groups are combinable into grammatical groups, phrases or clauses. This is a task of bridging a sequence of words input and the grammatical description of *instantial syntagmatic organizations* involving *configurations of grammatical functions.* In a second stage these grammatical functions can serve as paradigmatic context for further traversing the system network and extend to the full set of systemic features. Moreover they will play a crucial role in restricting and organising the search space for relevant and applicable network parts during analysis task.

Addressing the gap of easily accessible syntagmatic account within SFG framework~~,~~ can be done by first, providing information about ~~what~~ grammatical function operate at each rank, second which grammatical functions can be filled by which classes of units and third by providing relative and absolute account of ordering within each unit structure. This sort of information can guide building of a constituency backbone structure. As a second stage, as mentioned ~~before,~~ the unit classes and grammatical functions can operate as "hooks" on system network to guide the traversal in the same way the paradigmatic context available in the generation process.

Alternatively the problem of structure construction can be outsources ~~as~~ parsing with other grammars especially ~~that~~ there has been a lot of progress recently. Then the problem changes into creating a transformation mechanism to obtain the SF constituency structure rather than build it from scratch. Starting the SFG parsing process from a simple syntactic tree reduces the computational complexity by providing a set of reliable selections within the system network.

The second stage of constituent enrichment by network traversal can be further aided by checking an arbitrary set of patterns for preselecting even more features recoverable via lexico-syntactic patterns. The pattern recognition plays an essential role in current parsing method for fleshing out the constituent backbone with systemic selections.

1.6 On theoretical compatibility and reuse

In the past decades ~~there have been made significant progresses~~ in natural language parsing framed in one or another linguistic theory each adopting a distinct perspective and set of assumptions about language. The theoretical layout and the available resources influence~~s~~ directly what ~~and how~~ is ~~being~~ implemented into the parser and each implementation approach encounters challenges that may or may not be common to other approaches in the same or other theories.

Par~~e~~ses for one theoretical framework may face common or different problems across theories, but as well as the solutions. The successes and achievements in any school of though ~~should~~ be regarded as valuable cross theoretical results to the degree ~~the~~ links and correspondences can be established. Therefore reusing components that have been ~~proved~~ to work and yield “good enough results” is a strong pragmatic motivation for deriving herein described parsing method.

Present work lays first some cross theoretical correspondences and then some inter-grammatical links. It demonstrates how selected grammatical frameworks namely *Systemic Functional Grammar*, *Dependency Grammar* and *Governance & Binding Theory* relate to each other and to which degree they are compatible to undergo a conversion process and to show that simple patterns carrying grammatical information can be used to enrich ~~syntactically and semantically~~ ~~the~~ parse structures. And here is a brief motivation for selecting these frameworks.

In the last years *Dependency Grammar* (Tensiere 2015) became quite popular in natural language processing world favoured over phrase structure grammars. The grammatical lightness and the the modern algorithms implemented into dependency parsers such as Stanford Dependency Parser (Marneffe et al. 2006), MaltParser (Nivre 2006), MSTParser (McDonald et al. 2006) and Enju (Miyao & Tsujii 2005) are increasingly efficient and highly accurate.

I employ Stanford Dependencies (Marneffe & Manning 2008b,a; Marneffe et al. 2014) as a starting point which provides information about functional dependencies between words and grants direct access to the predicate-argument relations which are not readily available ~~form~~ the phrase structure parser and can be used off the shelf for real world applications. Regardless of being a simple grammatical framework which accounts for the syntactic relations between words, Stanford dependency grammar is structurally and functionally compatible to SFG. The account it provides for the word dependencies can be viewed also in functional terms and I expand this idea in Chapter 3. It is a much more suitable foundation for building the SFG syntactic structure than phrase-structure trees, as well as for making more delicate grammatical distinctions (a process highlighted in Section 1.9 and explained in detail in Chapter 6).

The current parsing process requires accounting for *null elements* which are not covered by the dependency grammar. As a solution I turned to a part of Chomsky's Transformational Grammar(Chomsky 1957), Government and Binding Theory (GBT) (Chomsky 1981; Haegeman 1991), to identify and create the Null Elements to support the semantic parsing. I introduce GBT and provide inter-grammatical links between towards dependency grammar in Chapter 4.

~~C~~urrent work is the first one to parsing with a dependency backbone, all previous ones using context-free grammars. There are other good candidates to serve as backbone (CCG, TAG, FCG, HPSG etc.) but a broad investigation of parsers and the compatibility of their linguistic theoretical frameworks to SFL is outside the scope of this thesis.

1.7 Previous works on parsing with Systemic Functional Grammars

There have been various attempts to ~~parsing~~ with SFGs. This section covers the most significant attempts to parse with a Systemic Functional Grammar. The firs attempt was made by Winograd (Winograd 1972) which was more than a parser, it was an interactive ~~a~~ natural language understanding system for manipulating geometric objects in a virtual world.

Starting from early 1980s onwards, Kay, Kasper, O'Donell and Bateman tried to parse with Nigel Grammar (Matthiessen 1985), a large and complex natural language generation (NLG) grammar for English used in Penman generation project. Other attempts by O'Donoghue (1991), Weerasinghe (1994), Souter (1996), Day (2007) aim for corpus based probability driven parsing within the framework of COMMUNAL project starting from late 1980s.

In a very different style, Honnibal (2004); Honnibal & Curran (2007) constructed a system to convert Penn Treebank into a corresponding SFGBank. This managed to provide a good conversion from phrase structure trees into systemic functional representation covering sentence mood and Thematic constituency (a kind of analysis in SFL which is not considered in current work). Transitivity has not been covered there because of its inherently semantic nature but it is in the current work.

1.7.1 Winograd's SHRDLU

SHRDLU ~~is~~ an interactive program for understanding ~~(if limited)~~ natural language written by Terry Winograd at MIT between 1968-1970. It carried a simple dialogue about a world of geometric objects in a virtual world. The human could as the system to manipulate objects of different colours and shapes and ~~the~~ ask questions about what ~~has~~ been done or the new state of the world.

~~It~~ is recognised as a landmark in natural language understanding demonstrating that a connection with artificial intelligence ~~is~~ possible if not solved. However, ~~his~~ success was not due to the use of SFG syntax but rather due to small sizes of every system component to achieve a fully functional dialogue system. Not only ~~it~~ was parsing the input but it was developing an interpretation of it, reason about it and generat~~e~~ appropriate natural language response.

Winograd combined the parsing and interpretation processes ~~such~~ that the semantic interpreter was actually guiding the parsing process. The knowledge of syntax was encoded in the procedures of interpretation program. He also implemented an ~~ingenious~~ backtracking mechanism where ~~the~~ the program ~~does~~ not simply go back, like other parsers, to try the next possible combination choice but actually takes a decision on what shall be tried next.

Having data embedded into the program procedures, as Winograd did, makes it non-scalable for example in accommodation of larger grammars and knowledge bodies and unmaintainable ~~on~~ the long term as it becomes increasingly difficult to make changes (Weerasinghe 1994).

1.7.2 Kasper

Bob Kasper ~~in 1985 being~~ involved in Penman generation project embarked on the mission of ~~testing~~ if the Nigel grammar, then the largest available generation grammar,

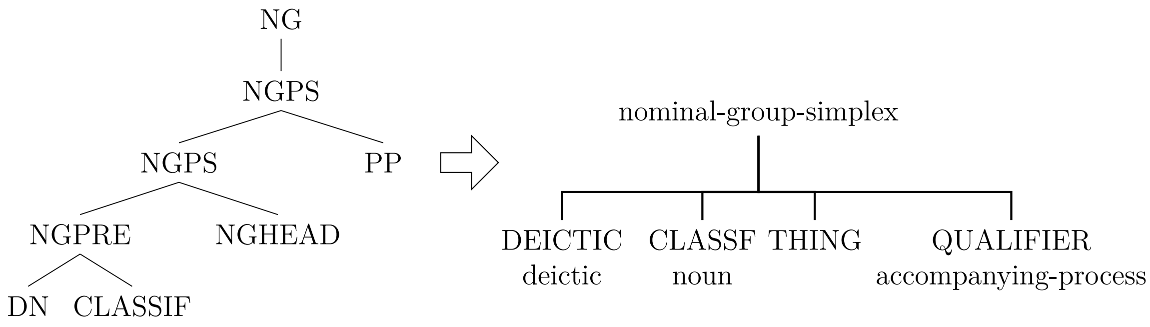


Fig. 1.3 Transformation from phrase structure into systemic constituency structure. Rule example from O'Donnell & Bateman (2005).

~~was suitable~~ for natural language parsing. Being familiar with Functional Unification Grammar (FUG), a formalism developed by Kay and tested in parsing (Kay 1985) which caught on popularity in computational linguistics regardless of Kay's dissatisfaction with results, Kasper decided to re-represent Nigel grammar into FUG.

Faced with tremendous computational complexity, Kasper (1988) decided to manually create the phrase-structure of the sentences with hand-written rules which were mapped onto a parallel systemic tree structure. Kasper in 1988 was the first one to parse with a context-free backbone. He first parsed each sentence with a Phrase Structure Grammar (PSG), typical to Chomsky's Generative Transformational Linguistics Chomsky (1957). He created a set of rules for mapping the phrase structure (PS) into a parallel systemic tree like the one depicted in Figure 1.3. When all possible systemic tree ~~were~~ created they were further enriched using information from Nigel Grammar (Matthiessen 1985).

Once the context-free phrase-structure ~~was~~ created using bottom-up chart parser it was further enriched from the FUG representation of Nigel grammar. This approach to parsing is called *parsing with a context-free backbone* as phrase-structure is conveyed as simplistic skeletal analysis, fleshed out by the detail rich systemic functional grammar.

Even though Kasper's system ~~is~~ represents the first attempt to parse with full Hallidayan grammar, ~~it's~~ importance is lowered, as O'Donnell & Bateman (2005) point out, by the reliance on phrase structure grammar.

1.7.3 O'Donnell

Since 1990, Mick O'Donnell experimented with several parsers for small Systemic grammars, but found difficulty when scaling up to larger grammars. While working in ~~EAD~~ project, funded by Fujitsu, he recompiled a subset of Nigel grammar into two resources: the set of possible function bundles allowed by the grammar (along with the bundles preselections) and a resource detailing which functions can follow a particular function (O'Donnell 1993, 1994).

This parser ~~was~~ operat~~ing~~ without a syntactic backbone directly from a reasonable scale SFG. However when scaled to the whole Nigel grammar the system became very slow because of the sheer size of the grammar and its inherent complexity introduced by multiple parallel classifications and functional combinations - a problem well described by Bateman (2008). Then O'Donnell wrote his own grammar of Mood that was more suitable for the parsing process and less complex than the recompiled Nigel.

In 2001, while working in a Belgian company O'Donnell came to conclusion that dependency grammars are very efficient for parsing. Together with two colleagues, he developed a simplified systemic grammar where elements were connected through a single function hence avoiding (functional) conflation. Also the ordering of elements was specified relative to the head rather than relative to each other.

More recently, O'Donnell in UAM Corpus Tool ~~embedded~~ a systemic chart parser (O'Donnell 2005) with a reduced systemic formalism. He classifies his parser as ~~a~~ left to right and bottom up with a custom lexicon where verbs are attributed features similar to Hallidayan process types and nouns a unique semantic category ~~like~~ thing-noun, event-noun, location-noun etc.

Because of previously reported complexity problems (O'Donnell 1993) with systemic grammars, the grammatical formalism is reduced to a singular functional layer of Mood-based syntactic structure (Subject, Predicate, Object etc.) ignoring the Transitivity (Actor/Goal, Sensor/Phenomenon etc.) and Textual (Theme/Rheme) analyses. O'Donnell ~~deals away with the~~ conflation except for the verbal group system network. He also employs a slot based ordering where elements do not relate to each other but rather to the group head only simplifying the number of rules and calculation complexity.

In ~~his paper~~ (O'Donnell 2005) ~~does not provide~~ a parser evaluation so its accuracy is ~~still~~ unknown ~~today.~~ The lexicon that was created is claimed to deal with word semantic classes but ~~it~~ is strongly syntactically based assigning a single sense to nouns and verbs ~~ignoring the peculiar aspect of language~~ polysemy. Moreover ~~it is not very clear~~ the framework within which the semantic classes have been generated.

1.7.4 O'Donoghue

O'Donoghue proposes a corpus based approach to parsing using *Vertical Strips* (O'Donoghue 1991). ~~They~~ are defined as a vertical path of nodes in a parse tree starting from the root down to the lexical items but not including those. ~~He~~ extracted the set of vertical strips from a corpus called Prototype Grammar Corpus together with their frequencies and probabilit~~y~~ of occurrence. This approach differ from the traditional one with respect to the kind of generalization it is concerned and specifically, the traditional approach~~ed~~ are oriented towards horizontal order while the vertical strip approach is concerned with vertical order in the parse tree.

To solve the order problem O'Donoghue uses a set of probabilistic collocation rules extracted from the same corpus indicating which strips can follow ~~a particular strip~~. He also created a lexical resource indicating for each word which elements can expand~~ed~~ it.

The parsing procedure is a simple lookup of words in the lexical resource selecting all possible elements it can expound and then selecting possible strips starting with the elements expounded by the word. Advancing from left to right for each sentence word more strips compatible with the previously selected ones ~~are~~ selected within the collocation network constrains. The parser finds all possible combinations of strips composing parse trees representing possible output parses.

The corpus from which the vertical strips were extracted ~~is~~ 100,000 sentences ~~large and was~~ generated with Fawcett's natural language generation system ~~and~~ was tested on the same corpus leaving unclear how ~~would~~ the parser behave on a real corpus. In 98% of cases the parser returns a set of trees (between 0 and 56) that included the correct one with an average of 6.6 trees per parse.

Actually, using a larger corpus could potentially lead to a combinatorial explosion in the step that looks for vertical strips. It would decrease the accuracy of the parse because of the higher number of possible trees per parse.

1.7.5 Honnibal

Honnibal (2004; 2007) describes how Penn Treebank ~~can~~ be converted into a SFG Treebank. Before assigning to parse tree nodes synthetic features such as mood, tense, voice and negation he first transforms the parse trees into a form that facilitates the feature extraction.

The scope of SFG corpus was limited to a few Mood and Textual systems leaving aside Transitivity because of its inherently lexico-semantic nature. He briefly describes how he structurally deals with verb groups, complexes and ellipses as functional structures are much flatter than those exhibited in the original Treebank. Then he describes how are identified metafunctional features of unit class, mood function, clause status, mood type, polarity, tense, voice and textual functions.

The drawback of his approach is that the Python script performing the transformation does not derive any grammar but rather implements directly these transformations as functions falling into the same class of problems like Winograd's SHRDLU. By doing so the program is non-scalable for example in accommodation of larger grammars and knowledge bodies and unmaintainable on the long term as it becomes increasingly difficult to make changes.

1.8 Current approach to parsing

The main problem in using SFGs for parsing is that they are very large and complex. Some parsing approaches use a syntactic backbone which is then flashed out with SFG description. ~~Other ones~~ use a reduced set or a single layer of SFG representation the ~~third ones~~ use an annotated corpus as the source of a probabilistic grammar. Regardless of ~~the~~ approach each limits the SFG in ~~a~~ one way or another balancing the depth of description with language coverage: that is either *deep description but a restricted language* or *shallow description but broad language coverage*.

Current approach is aligned with works of Honnibal, Kasper and O'Donnell with respect to using a backbone structure and enriching it with syntactic and semantic features. It relies on parse structures produces ~~in~~ other grammars and then translated to systemic functional constituency structur The contributions on theoretical compatibility and inter-grammatical transformations are ~~briefed~~ in ~~the~~ Section 1.6 ~~coming up next.~~

~~Current method~~ employs rules for graph traversal in order to build a parallel backbone constituency tree and rules for graph matching to enrich it with systemic features. This aims at keeping the language coverage broad at the expense of higher systemic delicacy. I cover basic Mood systems which is ~~way~~ less than what is available in Nigel grammar and leave it for ~~the~~ future work to ~~extend~~ the grammatical delicacy.

Nonetheless I attempt to cover some the lexico-semantic features as well. Parsing ~~Transitivity~~ ~~system, a task similar to~~ Semantic Role Labelling, requires large lexicogrammatical resource describing verb meanings in terms of their process type and participant roles. The semantically-oriented decomposition of clauses offered by SFL is still sufficiently closely tied to observable grammatical distinctions as to offer a powerful bridge to automatic analysis. Such descriptions are analogous to frame representations (Fillmore 1985) as found in FrameNet (Baker et al. 1998) or VerbNet (Kipper et al. 2008) applied in Semantic Role Labelling Task (Carreras & Màrquez 2005).

O'Donnell approaches this task by providing possible process types directly to the verb by employing self constructed lexicon where each word has syntactic and semantic features. Only recently a resource comparable to FrameNet and VerbNet has been produced in the SFL framework called Process Type Database (PTDB) (Neale 2002). PTDB which provides for each verb a process configuration (similar to ~~a~~ semantic frames) in terms of process type and participant roles. ~~Current work~~ represents the first attempt of using the PTDB to produce semantic (or Transitivity) analysis ~~which~~ combined with pattern patching method has ~~an~~ advantage over O'Donnell's parser. It enables to simultaneously assign, if matched, ~~the~~ systemic features to all clause constituents or not at all.

Another major advantage, as compared to Honnibal's approach is that the grammar and the program are carefully disconnected so that the code is maintainable and scalable with ~~the~~ respect to size of the grammar. This makes it possible to choose rather pragmatically which graph patterns to consider for parsing depending on the scope of task at hand.

1.9 The parsing process overview

Te parser follows a pipeline architecture depicted in Figure 1.4 where staring from an input text ~~gradually~~ a rich systemic functional constituency structure is built. This section~~s~~ provides an overview to the ~~building~~ process.

In the figure there are three types of boxes. The rounded rectangles represent the parsing steps. They linearly flow from one to the next ~~one~~ via green trapezoids boxes, on the left side, which represent intermediary data. On the right side are positioned double edged orange trapezoids representing ~~some~~ fixed resources used as additional input for some steps. For example *constituency graph creation* step takes a normalised dependency graph for input and produces a constituency graph as output.

The entire process starts with some input English text and ends with production of Rich Constituency Graph. The Input text is first parsed with a dependency parser. For the current work Stanford dependency parser was chosen for its dependency relations, parse accuracy and the continuous efforts put into its development ~~(motivated in Section 1.6)~~.

The dependency graphs often contain errors ~~some~~ of ~~which~~ are predictable, easy to identify and correct. ~~Also~~ some linguistic phenomena are treated in a slightly different manner than proposed in the current thesis. Therefore the dependency graph

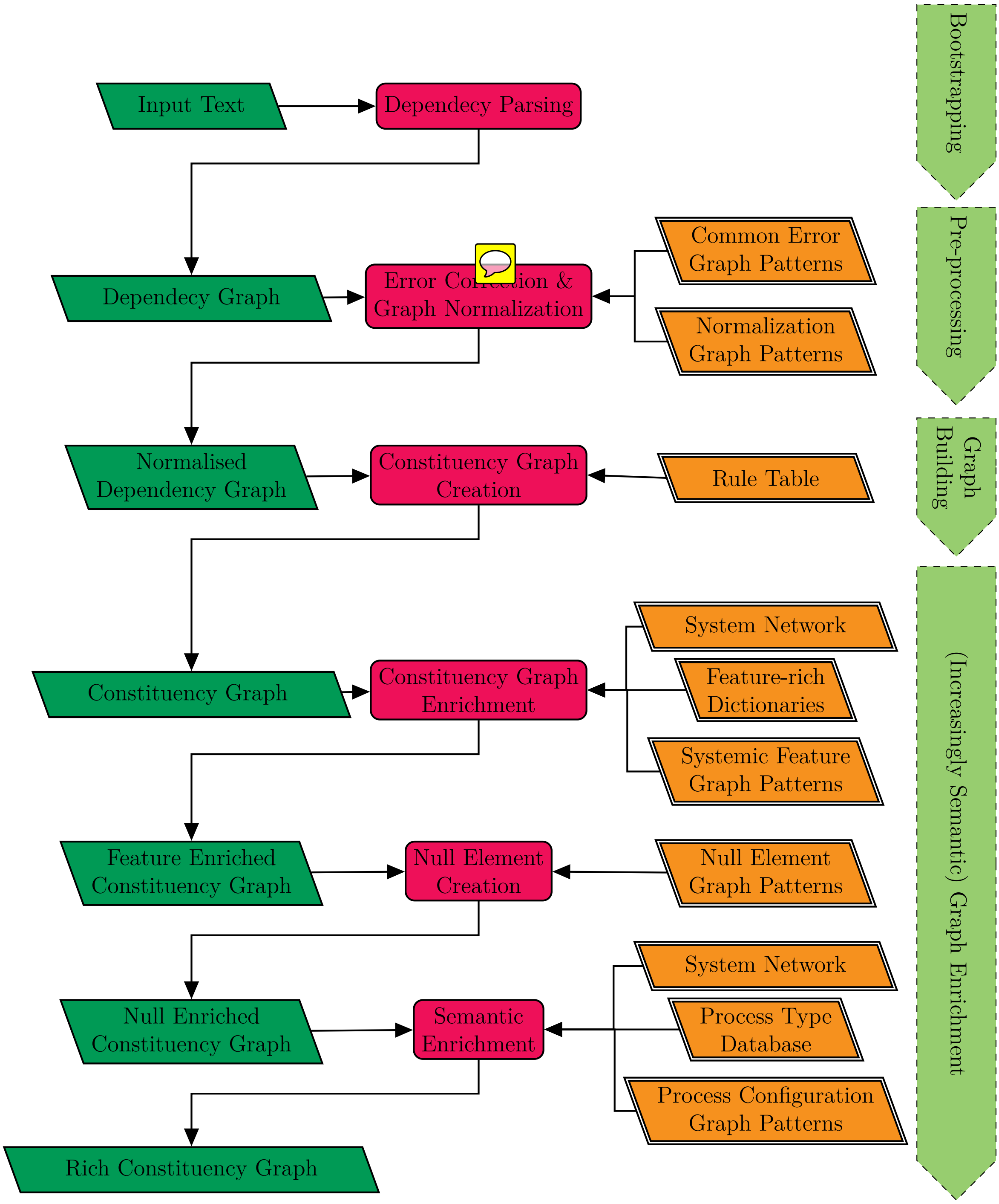


Fig. 1.4 The parsing process pipeline

produced by Stanford parser is *corrected and normalised* using pattern matching against ~~a~~ collections of known errors and ~~one~~ of normalization rules.

Once normalised the dependency graph is ready to guide the *building process* of the systemic functional constituency graph. It represents, in a way, transformation of the dependency graph, and serves a syntactic backbone on which the subsequent enrichment phases are performed.

Next follow two phases where the syntactic backbone is *enriched* with features some of which ~~bear~~ a *syntactic* ~~whereas~~ other a *semantic* nature. In between these enrichment phases there is a construction process which produces structural changes to the backbone adding some *empty constituents* that play a role in semantic enrichment. The enrichment phases use additional resources such as *system networks*, *feature rich lexicons*, *graph patterns* and *semantic databases*. The *null element creation* process also needs a collection of graph patterns for identifying where and what kind of null elements occur. The final result of the process is a *Rich Constituency Graph* of the original text comprising a ~~plenitude of~~ systemic feature selections associated with constituting units of structure.

1.10 Research questions and contributions

This thesis addresses the following questions:

- What is a computationally feasible method to parse with systemic functional grammars with a syntactic backbone?
- To what degree are Stanford Dependencies suitable as a syntactic backbone for Systemic Functional Grammar parsing?
- How can Process Type Database be used as a resource for SFG Transitivity parsing?
- How can Government and Binding Theory be used for detecting external predicate arguments in the context of SFG Transitivity parsing with PTDB?

Also it brings the following contributions:

- The analysis of theoretical and practical compatibility between the syntactic structures of Stanford Dependency and Systemic Functional Grammars along with an implemented method to transform from one structure to another.
- A fast engine for graph pattern matching which can also update and insert new nodes.

- A flexible and expressive method to represent systemic features as graph patterns together with two strategies for choice propagation in the systemic networks.
- A set of pattern graphs covering Mood, Transitivity and other smaller system networks.
- A method to transform PTDB into a set of Transitivity graph patterns.
- Derived principles and generalizations from the Government and Binding Theory (GBT) and represented them as graph patterns used to identify the covert elements of the clause that are explicitly mentioned outside the clause borders. These generalizations serve for semantic parsing where is very helpful to identify the external arguments of verbs.
- Development of a test corpus and evaluation of the parser.

1.11 Thesis organisation

The remaining of this thesis is organised as follows.

Chapter 2 ~~explains in~~ parallel Cardiff and Sydney theories ~~of grammar followed by a discussion of structure units of each grammar~~. When juxtaposed, weaknesses and strengths of each school emerge in contrast to each other ~~on~~ aspects ~~like~~ *structure, dependency relations, unit classes, systemic networks, rank scale and unit complexing.* I use elements of both grammars ~~therefore~~ I explain my stance on each of the above issues and ~~argument~~ the choices. In similar manner I discuss the syntactic and semantic units of each grammar ~~even if the systemic functional linguistics does not make such distinction it is useful in~~ establish~~ing~~ links to mainstream ~~methods~~ ~~for~~ language processing. ~~Basically,~~ this chapter presents the mixed grammar and its theoretical underpinning through the comparative discussion between two schools in SFL.

Chapters 3 and 4 introduce *Dependency Grammar* and *Governance and Binding Theory (GBT)*. Both frameworks are used as departing points to build the SFG structure. The cross-theoretical correspondences together with specific inter-grammatical links are developed in the same chapters.

Chapter 5 formally defines the structures used in this thesis and the operations on ~~them.~~ Important to mention structures are *feature rich graphs*, *ordered conjunction sets*, *feature structures* and *system networks*; whereas important operations are the varieties of *graph matching* and *pattern graph matching.*

Chapters 6 and 7 explain how the parsing process evolves starting from the dependency graph towards a constituency graph and then towards increasingly semantic constituency graph through its feature features. This suite of algorithms and the pipeline has a Python implementation called Parsimonious Vole[1]. A limited empirical evaluation of the parser is provided in Chapter 8. It describes the evaluation methodology, the gold standard used and highlights strengths and weaknesses of the current implementation.

~~The last part of the~~ thesis sets future directions explore and concludes on the current work (Chapter 9).

[1] Parsimonious Vole: https://bitbucket.org/lps/parsimonious-vole

Chapter 2

The systemic functional theory of grammar

There are two variants of Systemic Functional Grammars: the *Sydney Grammar* started in 1961 by Halliday (2002) and the *Cardiff Grammar* proposed by Fawcett (2008) which is a simplification and an extension of Sydney Grammar.

To understand the underlying motives and how exactly they are different we shall start looking at the theories of grammar before we look at the grammars proposed in Sydney and Cardiff SFL schools.

What is the difference between grammar and the theory of grammar, you may ask. If the grammar describes language in terms of categories, functions, relations, structures, meaning choices etc. then the theory of grammar defines what are the concepts that should be used to describer a grammar i.e. categories, functions, relations are and how they related to one another. Having a solid theory of grammar contributes to explaining what language is and how it works. It also frames how language is analysed by either human or machines.

This chapter discusses comparatively Halliday's (Halliday 2002) and Fawcett's (Fawcett 2000) ~~theoretical foundations~~ of SFL.

2.1 A word on wording

Before going into deeper discussion I ~~would like to~~ first make ~~a few~~ terminological clarifications on the terms: grammar, grammatics, syntax, semantics and lexicogrammar. I start with ~~a few~~ definitions adopted in ~~the~~ "mainstream" generative linguistics and then present how the same terms are discussed in systemic functional linguistics.

~~A.~~ Radford, a generative linguist, in the "Minimalist Introduction to Syntax" (1997), starts with a description of grammar as a field of study, which, in his words is traditionally subdivided into two inter-related areas of study: syntax and morphology.

Definition 2.1.1 (Morphology (Radford))**.** Morphology is the study of how words are formed out of smaller units (traditionally called morphemes) (Radford 1997: p.1).

Definition 2.1.2 (Syntax (Radford))**.** Syntax is the study of how words can be combined together to form phrases and sentences. (Radford 1997: p.1)

Halliday, in the context of rank scale discussion (Halliday 2002: p. 51), refers to the traditional meaning of syntax as the *grammar above the word* and to morphology as *grammar below the word*. Such distinction, he stresses, has no theoretical status. His precursor, Firth, puts it in following terms: "[...] the distinction between morphology and syntax is no longer useful or convenient in descriptive linguistics." (Firth 1957: p.14)

Radford adds that, traditionally, grammar is not only concerned with the principles governing formation of words, phrases and sentences but also with principles governing their interpretation. Therefore *structural aspects of meaning* are said to be also a part of grammar.

Definition 2.1.3 (Grammar (Radford))**.** [Grammar is] the study of the principles which govern the formation and interpretation of words, phrases and sentences. (Radford 1997: p.1)

Interestingly enough, the Definition 2.1.3 makes not mention at all to the lexicon. This is because the formal grammars focus primarily on unit classes and how they are accommodated in various structures and so in formal linguistics the lexicon is disconnected from the grammar. The systemic grammar, on the other hand, along with formal descriptions of grammatical categories and structures, include lexicon as part of grammar to form a *lexicogrammar*.

Another important aspect to notice is that the grammar is defined as a field of study rather than a set of rules. Halliday, since his early papers, became conscious of the confusion made in the literature between a study of a phenomenon with the phenomenon itself. By analogy to language as phenomenon and linguistics as the study of the phenomenon, Halliday adopts the same wording for **grammar** as phenomenon and *grammatics* as the study of grammar; the same distinction holds for *syntax* and *syntactics*.

Definition 2.1.4 (Grammatics (Halliday))**.** Grammatics is a theory for explaining grammar (Halliday 2002: p.369)

~~E.~~ Moravcsik, another generative linguist stresses the same distinction , in her “An introduction to syntax” (Moravcsik 2006), and presents two ways in which the word “syntax” is used in literature: (a) in reference to a particular aspect of grammatical structure and (b) in reference to a sub-field of descriptive linguistics that describes this aspect of grammar.

In her words “syntax describes the selection and order of words that make well-formed sentences and it does so in as general a manner as possible so as to bring out similarities among different sentences of the same language and different languages and render them explainable. [...] syntax rules also need to account for the relationship between string of word meanings and the entire sentence meaning, on one hand, and relationship between strings of word forms and the entire sentential phonetic form, on the other hand.” (Moravcsik 2006: p.25)

In her definition of grammar she includes lexicon and semantics which is somewhat more explicit statement than Radford’s “interpretation”. She is getting, in Definition 2.1.5, somewhat closer to what grammar stands for in SFL - Definition 2.1.6.

Definition 2.1.5 (Grammar (Moravcsik))**.** ... maximally general analytic descriptions, provided by descriptive linguistics, [are] called grammars. A grammar has five components: phonology (or, depending on the medium, its correspondent e.g. morphology), lexicon, syntax and semantics(Moravcsik 2006: pp.24–25).

Definition 2.1.6 (Grammar (Halliday))**.** To Halliday, lexico-grammar or short grammar is a part of language and it means the wording system - the “lexical-grammatical stratum of natural language as traditionally understood, comprising its syntax, vocabulary together with any morphology the language may display [...]” (Halliday 2002: p.369).

The last point I want to mention is the approach to semantics. Formal grammars aim to account for the realisation variations, that is formation of words, phrases and sentences along with their arrangements and ~~the~~ mention~~s~~ of semantics ~~merely refer to~~ the “formal aspect of meaning”.

By contrast, a systemic grammar is a functional grammar which means (among other things) that it is semantically motivated, i.e. “natural”. So the fundamental distinctions between formal and functional grammars is the semantic basis for explanations of structure.

Also, in SFL, the meaning is being approached from a semiotic perspective, ~~punting~~ the linguistic semantics in perspective with the linguistic expression and the real world situation.

In this respect, Lemke (1993) offers a well formulated theoretical foundation that "human communities are eco-social systems that persist in time through ongoing exchange with their environment; and the same holds true for any of their sub-subsystems [...]" including language. The social practices constituting such systems are both material and semiotic, with a constant dynamic interplay between the two. (Halliday 2002: ~~p.~~387)

The term *semiotic* means to Halliday oriented towards meaning rather than sign. In other words, the interaction is between *the practice of doing and the practice of meaning*.

As the two sets of practices are strongly coupled, Lemke points out that there is a high degree of redundancy between the material-semiotic interplay. And it perfectly resonates with Firth's idea of *mutual expectancy* between the text and the situation.

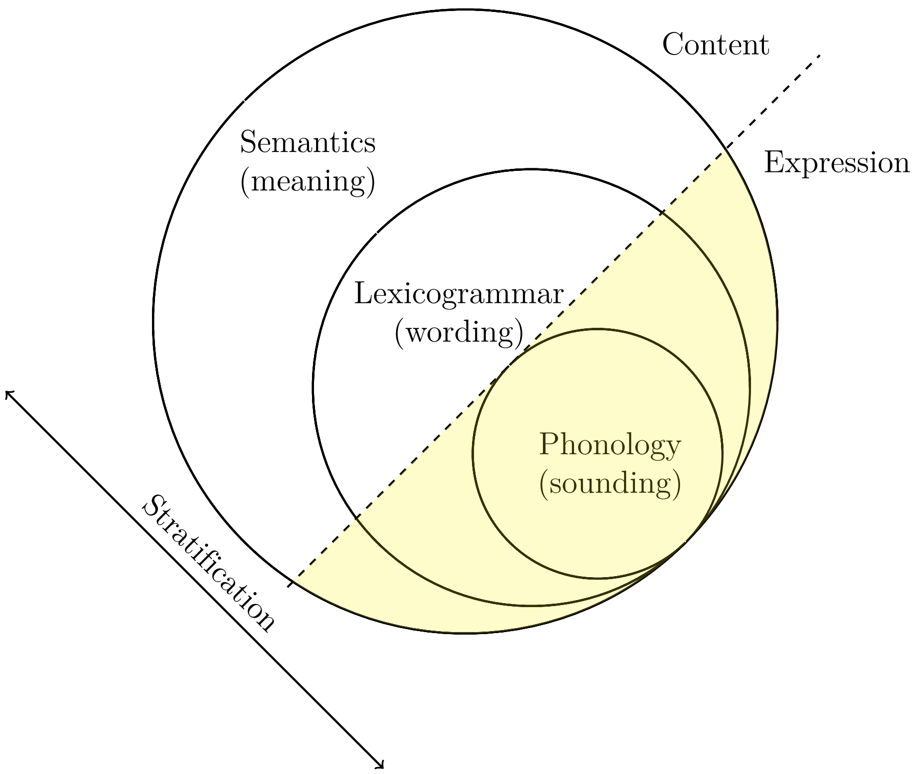


Fig. 2.1 The levels of abstraction along the realisation axis

Having that said, the best way to relate the formal and the systemic functional grammars is by placing the two of them along the stratification axis.

The SFL model defines language as a resource organised into three strata: phonology (sounding), lexicogrammar (wording) and semantics (meaning). Each is defined according to its level of abstraction on the realisation axis. The realisation axis is divided into two planes: the expression and the content planes. The first strat~~a~~ (i.e. phonology) belongs to the *expression plane* and the last two (lexicogrammar and semantics) belong to the *content plane*. But the division is not so clear because some parts of semantics and lexicogrammar transcend from content into expression plane. In this context, the formal grammar could be localised entirely within the expression plane, including the phonology/morphology, syntax, lexicon and semantics but stripped of any explanations in terms of the meaning potential available in the content plane.

2.2 Sydney theory of grammar

Halliday (2002) proposes four fundamental categories of grammar: *unit*, *structure*, *class* and *system*. Each of these categories is logically derivable from and related to other ones in a way that they mutually define each other. These categories relate to each other on three scales of abstraction: *rank*, *exponence*, *delicacy*. Halliday also uses three scale types: *hierarchy*, *taxonomy* and *cline*.

Definition 2.2.1 (Hierarchy)**.** Hierarchy [is] a system of terms related along a single dimension which involves ~~a~~ some sort of logical precedence. (Halliday 2002: p.42).

Definition 2.2.2 (Taxonomy)**.** Taxonomy [is] a type of hierarchy with two characteristics:

1. the relation between ~~a~~ terms and the immediately following and preceding one is constant
2. the degree is significant and is defined by the place in the order of a term relative to following and preceding terms. (Halliday 2002: p.42)

Definition 2.2.3 (Cline)**.** Cline [is] a hierarchy that instead of being made of a number of discrete terms, is a continuum carrying potentially infinite gradations. (Halliday 2002: p.42).

Next I define and introduce each category of grammar and the related concepts that constitute the theoretical foundation for the Sydney Theory of grammar.

2.2.1 Unit

Language is patterned activity of meaningful organization. The patterned organization of substance (*graphic* or *phonic*) along a linear progression is called *syntagmatic order* (or simply *order*).

Definition 2.2.4 (Unit)**.** The unit is a grammatical category that accounts for the stretches that carry grammatical patterns~~.~~ (Halliday 2002: p.42). The units carry a fundamental *class* distinction and should be fully identifiable in description~~.~~ (Halliday 2002: p.45).

Generalization 2.2.1 (Constituency principles)**.** The five principles of constituency in lexicogrammar are:

1. There is a scale or rank in the grammar of every language. That of English (typical of many) can be represented as: clause, group/phrase, word, morpheme.

2. Each unit consists of *one or more* units of rank next below.

3. Units of every rank may form complexes.

4. There is potential for rank shift, whereby a unit of one rank my be downranked to function in a structure of a unit of its own rank or of a rank below.

5. Under certain circumstances it is possible for one unit to be ~~enclsoed~~ within another, not as a constituent but simply in such a way as to split the other ~~oe~~ into two discrete parts. (Halliday & Matthiessen 2013: pp.9–10)

The relation between units is that of consistency for ~~witch~~ we say that a unit *consists of* other units. The scale on which the units are ranged is the *rank scale*. The rank scale is a levelling system of units supporting unit composition regulating how units are organised at different granularity levels from clause, to groups/phrases to words and the units of a higher rank scale consist of units of the rank next below. ~~The~~ Table 2.1 presents a schematic representation of the rank scale and its derived complexes.

Generalization 2.2.2 (Rank scale constraints)**.** The rank relations are constrained as follows:

1. downward *rankshift* is allowed i.e. the transfer of a given unit to a lower rank.

2. upward rankshift is not allowed.

Rank scale ↓	**Complexing**
	Clause complex
Clause	
	Group(/phrase) complex
Group(/phrase)	
	Word complex
Word	
	(Morpheme complex)
(Morpheme)	

Table 2.1 Rank scale of the (English) lexicogrammatical constituency

3. only whole units can enter into higher units~~.~~(Halliday 2002: p.44).

~~The~~ Generalization 2.2.2 taken as a whole means that a unit can include, in what it consists of, a unit of rank higher than or equal to itself but not a unit of rank more than one degree lower than itself; and not in any case a part of any unit~~.~~ (Halliday 2002: p.42).

2.2.2 Structure

Definition 2.2.5 (Structure)**.** The structure (of a given unit) is the arrangement of *elements* that take places distinguished by order relationship (Halliday 2002: p.46).

Definition 2.2.6 (Element)**.** ~~The~~ element is defined by the place stated as absolute or relative position in sequence and with ~~the~~ reference to the unit next below (Halliday 2002: p.47).

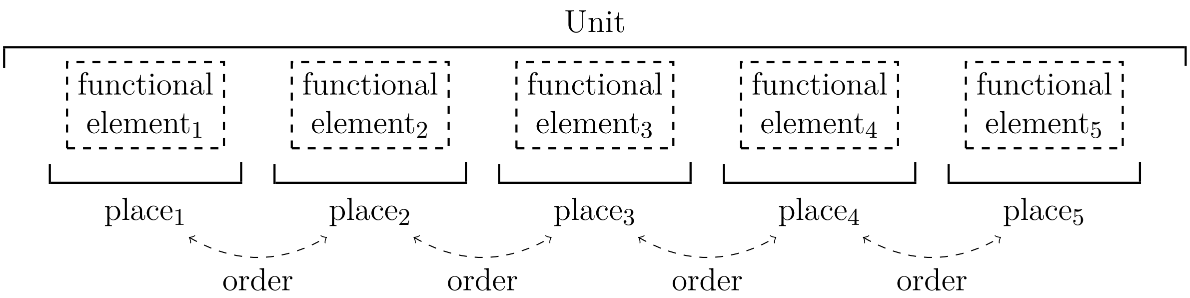


Fig. 2.2 The graphic representation of (unit) structure

We say that a~~n~~ unit is composed of elements located in places and that its internal structure is accounted via elements in terms of functions and places taken by the lower (constituting) units or lexical items. The graphic representation of the unit structure is depicted in Figure 2.2. The unit structure is referred in linguistic terminology as *constituency* (whose principles are enumerated in Generalization 2.2.1). In the unit structure, the elements resemble an array of empty slots that are *filled* by other units or lexical items.

2.2.3 Class

Definition 2.2.7 (Class)**.** The class is that grouping of members of a given unit which is defined by operation in the structure of the unit next above (Halliday 2002: p.49).

Halliday defines class (Definition 2.2.7) as likeness of the same rank "phenomena" to occur together in the structure. He adopts a top-down approach stating that the class of a unit is determined by the *function* (Definition 2.2.9) it plays in the unit above and not by its internal structure of elements. In SG the structure of each class is ~~well~~ accounted in terms of syntactic variation recognizing six unit classes: *clause*, *prepositional phrase* and *nominal*, *verbal*, *adverbial* and *conjunction* groups. Sydney grammar is briefly summarised in the Appendix 9.6.

2.2.4 System

Structure is a syntagmatic ordering in language capturing regularities and patterns which can be paraphrased as *what goes together with what*. However, language is best represented as a set of system networks (Definition 2.2.8) which is a paradigmatic ordering in language describing *what could go instead of what* (Halliday & Matthiessen 2013: p. 22).

This is an essential assumption of systemicists is that the language is best represented in the form of system networks and not as an inventory of structures. The structure of course is a part of language description but ~~it is~~ only a syntagmatic manifestation of the systemic choices (Halliday & Matthiessen 2013: p.23).

Definition 2.2.8 (System)**.** A system is a set of mutually exclusive set of terms referring to meaning potentials in language and are mutually defining. It always means a *closed system* and has the following characteristics:

1. the number of terms is finite,

2. each term is exclusive of all others,

3. if a new term is added to the system it changes the meaning of all other terms. (Halliday 2002: p.41)

A class is a grouping of items identified by operation in the structure. It is not a list of formal items but an abstraction from them. By increase in *delicacy* the class is broken into secondary classes which stand in the relation of exponent to an element of primary structure of the unit next above. This breakdown gives a system of classes that constitute choices implied by the nature of the class~~.~~ (Halliday 2002: p.41)

2.2.5 Functions and metafunction

Definition 2.2.9 (Function)**.** The functional categories or functions provide an interpretation of grammatical structure in terms of the overall meaning potential of the language~~.~~ (Halliday & Matthiessen 2013: p.76).

Most constituents of ~~the~~ clause structure, however, have more than one function which is called a *conflation of elements*. For example in the sentence “Bill gave Dolly a rose”, “Bill” is the Actor doing the act of giving but also the Subject of the sentence. So we say that Actor and Subject functions are conflated in the constituent “Bill”. This is ~~exactly the point~~ where the concept of *metafunction* or *strand of meaning* comes into the picture. The Subject function is said to belong to the *interpersonal metafunction* while Actor function belongs ~~in~~ the *experiential metafunction*.

Halliday identifies three fundamental dimensions of structure in the clause each ~~with distinct~~ meaning: *experiential*, *interpersonal* and *textual*. He refers to them as *metafunctions* and they account of how language meaning has evolved. Table 2.2 presents metafunctions and their reflexes in ~~the~~ grammar as proposed in (Halliday & Matthiessen 2013: p.85).

Metafunction	**Definition (kind of meaning)**	**Corresponding status in clause**	**Favored type of structure**
experiential	construing a model of experience	clause as representation	segmental (based on constituency)
interpresonal	enacting social relationship	clause as exchange	prosodic
textual	creating relevance to context	clause as message	culminative
logical	constructing logical relations	-	iterative

Table 2.2 Metafunctions and their reflexes in the grammar

Generalization 2.2.3 (Exhaustiveness principle)**.** Everything in the wording has some function at every rank but not everything has a function in every dimension of structure~~.~~ (Halliday 2002; Halliday & Matthiessen 2013)

With respect to structure and metafunctions, Halliday formulates the general principle of *exhaustiveness* (Generalization 2.2.3) saying that clause constituents have at least one and may have multiple functions in different strands of meaning~~,~~ however ~~it~~ does not mean that it must have a function in each of them.

This principle implicitly relates to the ~~economic~~ property of language meaning that it naturally evolves towards the shortest and most effective way of expressing a meaning. There is nothing meaningless thus every piece of language must be explained and accounted for in the lexicogrammar.

2.2.6 Lexis and lexicogrammar

In SFL the terms *word* and *lexical item* are not really synonymous. They are ~~strongly~~ related but they refer to different things. The term "word" is reserved (in early Halliday) for the grammatical unit of the lowest rank whose *exponents* are lexical items.

Definition 2.2.10 (Lexical Item)**.** In English, a lexical item may be a *morpheme*, *word* (in traditional sense) or *group (of words)* and it is assigned to no rank~~.~~ (Halliday 2002: p.60)

Examples of lexical items are ~~all of~~ the following ~~ones:~~ " 's " (the possessive morpheme), "house, walk, on" (words in traditional sense) and "in front of, according to, ask around, add up to, break down" (multi word prepositions and phrasal verbs)

If most linguists treat the grammar and lexis as discrete phenomena, Halliday brings them together as opposite poles of the same cline. We say that they are paradigmatically related through delicacy relation. He refers to this merge as *lexicogrammar* and he expressed his dream that one day linguists will be able to turn whole linguistic form into (lexico)grammar showing that lexis is the most delicate grammar.

Hasan (2014), explores ~~the reality of Halliday's dream in terms of project~~ feasibility ~~and exploring the implications~~ of what would it mean to turn the "whole linguistic form into grammar". This then implies two completely new assumptions: that lexis is not form and that its relation to semantics is unique (challenging the problems of polysemy). It would be the function of the lexicogrammar to map the multiple *meta-functional strata* into a unified structure.

2.3 Cardiff Theory of grammar

This section present the theory of grammar as conceived by Robin Fawcett at University of Cardiff. The biggest difference to Hallidayan theory is renouncing the concept of rank scale which has an impact on the whole theory. As a consequence, to accommodate the lack of rank-scale, Fawcett adapts the definitions of the fundamental concepts and slightly changes the choice of words.

In 2000 Robin Fawcett present~~s~~ a theory of grammar in contrast to some aspects of Michael Halliday's grammar 2002. One of the main differences is the rejection of the rank scale concept. Another is the bottom-up approach to unit definition as opposed to top-down one advocated by Halliday. These two and few other ~~discrepancies~~ have ~~quite an~~ important implication ~~on~~ the overall theory of grammar and ~~of course~~ the grammar itself.

Fawcett (2000) proposes three fundamental categories in the theory of grammar: *class of unit*, *element of structure* and *item*. Constituency is a relation accounting for prominent compositional dimension of language. However a unit does not function directly as a constituent of another unit but via a specialised relation~~,~~ Fawcett breaks down ~~constituency~~ into three relations: *componence*, *filling* and *exponence*. Informally is said that a unit is composed of elements which are either filled by another unit or expounded by an item. He also proposes three secondary relations of *coordination*, *embedding* and *reiteration* to account for ~~the full~~ range of syntactic phenomena.

2.3.1 Class of units

Definition 2.3.1 (Class of Unit)**.** The class of unit [...] expresses a specific array of meanings that are associated with each one of the major classes of entities in semantics [...and] are to be identified by the elements of their internal structure (Fawcett 2000: p.195).

Class of unit is determined based on ~~its~~ internal structure i.e. by its elements of structure (and not by the function it plays in the parent unit).

Fawcett takes a semantic stance in classifying units ~~which~~ in line with Saussurean approach to language. He proposes that in English there are four major semantic classes of entities: situations, things, qualities (of situations and things) and quantities (typically of things but also of situations and qualities) corresponding to major syntactic units of *clause*, *nominal group*, *prepositional group*, *quality group* and *quantity group* (Fawcett 2000: p. 193–194) along with a set of minor classes such as *genitive cluster* and *proper name cluster*.

~~His~~ classification is based on the idea that the syntactic and semantic units are mutually determined and supported by grammatical patterns. However those patterns ~~are~~ beyond the syntactic variations of the grammar and blend into lexical semantics.

2.3.2 Element of Structure

Definition 2.3.2 (Element of Structure)**.** The elements of structure are immediate components of classes of units and are defined in terms of their *function* in expressing meaning and not in terms of their absolute or relative position in the unit~~.~~ (Fawcett 2000: pp.213–214).

Generalization 2.3.1. Definition 2.3.2 leads to the following two principles:

1. Every element in a given class of unit serves a function in that unit different from the function of the sibling elements.
2. Every element in every class of unit will be different from every element in every other class of unit. (Fawcett 2000: p.214)

The elements (of structure) are functional slots which define the internal structure of a~~n~~ unit but still they are *located* in *places*. One more category that intervenes between element and unit is the concept of *place* which become essential for the generative versions of grammar.

There are two ways to approach place definition. The first~~,~~ is to treat places ~~a~~ positions of elements relative to each other (usually previous). This leads to the need ~~of~~ an *anchor* or a *pivotal element* which may not always be present/realised.

The second~~,~~ is to treat places as a linear sequence of locations at which elements may be located, identified by numbers "place 1", "place 2" etc. This place assignment approach is absolute within the unit structure and makes elements independent of each other. This approach has been used in COMMUNAL and Penman projects.

2.3.3 Item

Definition 2.3.3 (Item)**.** The item is a lexical manifestation of meaning outside syntax corresponding to both words (in the traditional sense), morphemes and either intonation or punctuation (depending whether the text is spoken or written)~~.~~ (Fawcett 2000: pp.226–232).

Items correspond to the leaves of syntactic trees and constitute the raw *phonetic* or *graphic* manifestation of language. The collection of items of a language is generally referred as *lexis*.

Since items and units are of different nature the relationship between an element and a (lexical) item must be different from that to a unit. We say that items *expound* elements and not that they *fill* elements as units do.

In traditional grammar *word classes* or *parts of speech* are a commonly accepted concept. However in SFL, it plays rather an orientation or an approximation role, precisely because the word classes do not properly correspond to the elements they expound. So terms as *noun* or *adjective* are useful to denote a class of words that expound a certain element of the structure, but such word class to element correspondence ~~shall~~ by no means treated as definite rule.

2.3.4 Componence

Definition 2.3.4 (Componence)**.** Componence is the part-whole relationship between a unit and the elements it is composed of. (Fawcett 2000: p.244).

Note that componence is not a relationship between a unit and its places; the latter, as discussed in Section 2.3.2, simply locationally relate elements of a unit to each other.

Componence intuitively implies a part-whole constituency relationship between the unit and its elements. But this is not the only view. Another perspective is the concept of *dependency* or strictly speaking *sister or sibling dependency* (because the traditional concept of dependency is parent-daughter relation). However the sister dependency is not necessary in the grammar model and is a by-product or second order concept that can be deduced from the constituency structure.

The (supposed) dependency relation between a modifier and the head, in the framework of SFG is, not a direct one ~~that form-centered linguists consider to be.~~ They simply assume that what modifier modifies is the head. Here however the general function of the modifiers is to contribute to the meaning of the whole unit which is anchored by the head. For example, in the nominal group, the modifier contributes to the description of the referent stated by the head. So the head realises one type of meaning that relates the referent while modifier realises another one. Both of them describe the referent via different kinds of meaning, therefore they are related indirectly to each other because the modifier does not modify the head but the referent denoted by the head. (Fawcett 2000: p.216)

Moreover the dependency relations are expressed between system networks and according to Fawcett this is the true place for dependencies in SFL.

2.3.5 Filling and the role of probabilities

Definition 2.3.5 (Filling)**.** Filling is the probabilistic relationship between a element and the unit lower in the tree that operates at that element~~.~~ (Fawcett 2000: p.238, 251).

Fawcett renounces the concept of rank scale and alternatively proposed the concept of *filling probabilities*. The probabilistic predictions are made in terms of filling relationship between a unit and an element of structure in a higher unit in the tree rather than being a relationship between units of different ranks. This ~~places~~ focus from the fact that a unit is for example a group, to what group class it is.

In this line of though, some elements of a clause are frequently filled by groups, but some other element ~~almost never being~~ rather expounded by items. The frequency varies greatly and is an important factor for predicting or recognizing either the unit class or the element type in the filling relationship.

Filling may add a single unit to the element of structure or it can introduce multiple coordinated units. Filling also makes possible the embedding relation. Both~~,~~ coordination and embedding relations make~~s~~ it possible to deal without inter-clausal *hypotaxis* and *parataxis* relations described in Sydney Grammar.

Note also that filling and componence are two complementary relations that occur in the syntactic tree down to the level when the analysis moves out of abstract syntactic categories to more concrete category of items via the relationship of exponence.

2.3.6 A few more concepts

Definition 2.3.6 (Exponence)**.** Exponence is the relation by which an element of structure is realised by a (lexical) item (Fawcett 2000: p.254).

Definition 2.3.7 (Coordination)**.** Coordination is the relation between units that fill the same element of structure (Fawcett 2000: p.263).

Coordination is usually marked by an overt *Linker* such as *and*, *or*, *but*, etc. and sometimes it is enforced by another linker that introduces the first unit such as *both*.

Coordination is through by Fawcett as being not between syntactic units but between mental referents. It always introduces more than one unit which are syntactically and semantically in similar (somehow) resulting in a *syntactic parallelism* which often leads to *ellipsis*.

Definition 2.3.8 (Reiteration)**.** Reiteration is the relation between successive occurrences of the same item expounding the same element of structure (Fawcett 2000: p.271).

Reiteration often is used to create the effect of emphasis such as for example "she's very very nice!". Like coordination, reiteration is a relation between entities that fill the same element of the unit structure which is problematic in my opinion and I further discuss it in Section 2.4.6.

Definition 2.3.9 (Embedding)**.** Embedding is the relation that occurs when a unit fills an element of the same class of units, i.e. when a unit of the same class occurs (immediately) above it in the tree structure (Fawcett 2000: p.264).

Fawcett ~~opens~~ embedding as a general principle as opposed to exceptional/controlled embedding indicated by Halliday. I will further discuss ~~it~~ in the context of rank-scale concept in Section 2.4.1

Definition 2.3.10 (Conflation)**.** Conflation is the relationship between two elements that are filled by the same unit having the meaning of "immediately after and fused with" and function as one element (Fawcett 2000: pp.249–250).

Conflation is useful in expressing multi-faceted nature of language when for example syntactic and semantic elements/functions are realised by the same unit for example the Subject and the Agent or Complement and Affected. Also conflation relations frequently occur between syntactic elements as well such as for example the Main Verb and Operator or Operator and Auxiliary Verb.

Definition 2.3.11 (Taxis)**.** *Taxis* represents the degree of interdependency between units systematically arranged in a linear sequence where *parataxis* means equal and *hypotaxis* means unequal status of units forming a *nexus* together.

Taxis play a central role in explaining the textual cohesion. Figure 2.3 depicts the Taxis and clause logico-semantic relations in a single system network. Here taxis relations are employed in explaining how clauses form a nexus with certain tactic relations to one another (Halliday & Matthiessen 2013: pp. 438 – 443). These concepts nevertheless are very useful at describing unit relations not only at the group and clause ranks but all the way down to smallest linguistic unit such as morphemes and

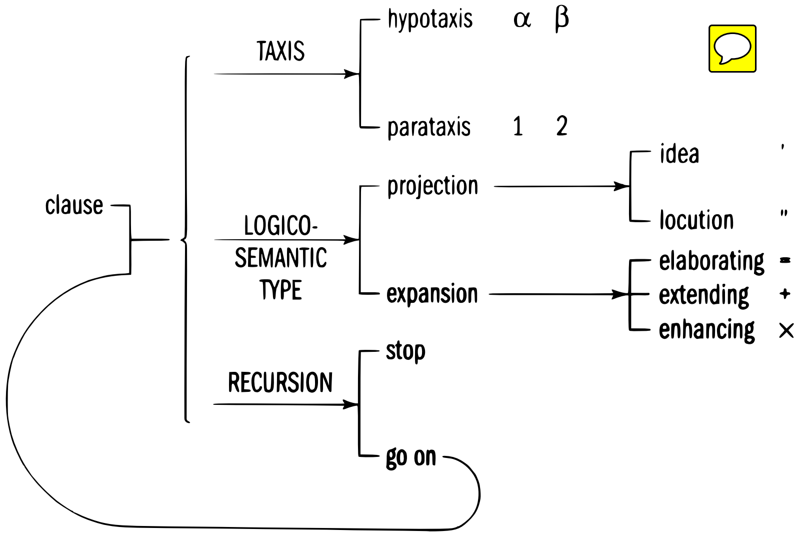


Fig. 2.3 Types of relations between clauses (Halliday & Matthiessen 2013: p.438)

phonemes. I will also employ the concept of taxis in the discussion of dependency relations in Section 3.5. Next I discuss the strengths and weaknesses of the two schools with ~~a~~ pragmatic goal of parsing in mind.

2.4 Critical discussion on both theories

The two sections above cover the definitions and fundamental concepts from each of the two systemic functional theories of grammar. ~~Current~~ ~~work~~ uses a mix of concepts from both theories and this section discusses in detail what ~~and why~~ is being adopted attempting a rather pragmatic reconciliation than a theoretical debate. Next I draw parallels and highlight correspondences between Sydney and Cardiff theories of grammar and where ~~needed~~ alter and present my position on the matter.

2.4.1 Relaxing the rank scale

The *rank scale* proposed by Halliday (2002) became ~~overtime~~ a highly controversial concept in linguistics. ~~W~~hether it is ~~a~~ suitable for grammatical description or not still continues ~~up to date~~. The historic development of this ~~polemic~~ is documented in (Fawcett 2000: p.309–338).

I consider rank scale a very useful dimension for unit classification and placement but the definition laid by the Sydney school is too rigid and thus I propose to relax it into a weaker version of it. The relaxation consists of dropping completely the *rank scale constraints* as enunciated in Generalization 2.2.2. An immediate consequence is that the *embedding* relation can be broadly defined as a naturally occurring phenomena in language at all ranks and not only for clauses as initially proposed.

Halliday's theory allows ~~the~~ downwards rank shift, forbids upwards rank shift and restricts the composition relation to engage only with whole units. Thus the unit may be composed of units of equal rank or a rank higher and cannot be composed of units that are more than one rank lower or parts of other units. The consequence of ~~above~~ is that each element of the clause is filled by a group which has its elements expounded by words.

(3) some very small wooden ones

The above rules often pose analysis difficulties and complications. For example in nominal groups what seems to be an element is not a single word but a group of words. Consider example 3 where Epithet *"very small"* is not a single word but a group (Halliday & Matthiessen 2013: pp. 390–396). This kind of phenomena introduce~~d~~ a *substructure* of modifiers and heads (see analysis in table 2.3) which complicates the general structure of the nominal group. Accordingly, the Epithet *"very small"* is composed of a head "small" and a modifier "very". This kind of intricate cases can be simplified through rank-shift constraints allowing the elements of a group to be filled by other groups or expounded by words.

some	**very**	**small**	**wooden**	**ones**
Deictic	*Epithet*		*Classifier*	*Thing*
Modifier				*Head*
	Sub-Modifier	*Sub-Head*		

Table 2.3 Sydney analysis of Example 3

An approach to describe units outside the rank-scale was suggested by Fawcett (2000) and Butler (1985). Because units are carriers of a grammatical pattern, they can be described in terms of their internal structure instead of their potential for operation in the unit above.

Fawcett proposes complete abandonment of rank scale replaces it with the filling probabilities to guide the unit composition simply mapping elements to a set of legal unit classes that may fill it. The above example, in CG the *"very small"* is analysed as a quality group that plays the role of Modifier (CG) in the nominal group as in table 2.4.

some	**very**	**small**	**wooden**	**ones**
Quantifying Determiner	*Modifier*		*Modifier*	*Head*
	Quality Group			
	Degree Tamperer	*Apex*		

Table 2.4 Cardiff analysis of Example 3

I maintain the idea of ranking the syntactic units because it is a pertinent classification with a clear correspondence to the types of meaning structures in the ideational space.

However I drop the constituency constraints and hence allowing the flexibility for elements to be filled by other units or in other words allow unit embedding. This approach removes the need of sub-structures in the unit elements reducing thus the structural complexity.

The rank system constrains had consequences on the phenomena of embedding defined by Halliday in Definition 2.4.1, which I consider way too restrictive.

Definition 2.4.1 (Embedding (strict))**.** Embedding is the mechanism whereby a clause or phrase comes to function as a constituent within the structure of a group, which is itself a constituent of a clause. (Halliday & Matthiessen 2013: p.242)

The weakening of constituency constrains makes embedding a normal (as defined in broad sense 2.3.9 by Fawcett) rather than an exceptional (as defined in a strict sense in 2.4.1 by Hallliday) phenomena. And I agree with Fawcett's definition because the human language allows construction of units that contain other units within them regardless of their class.

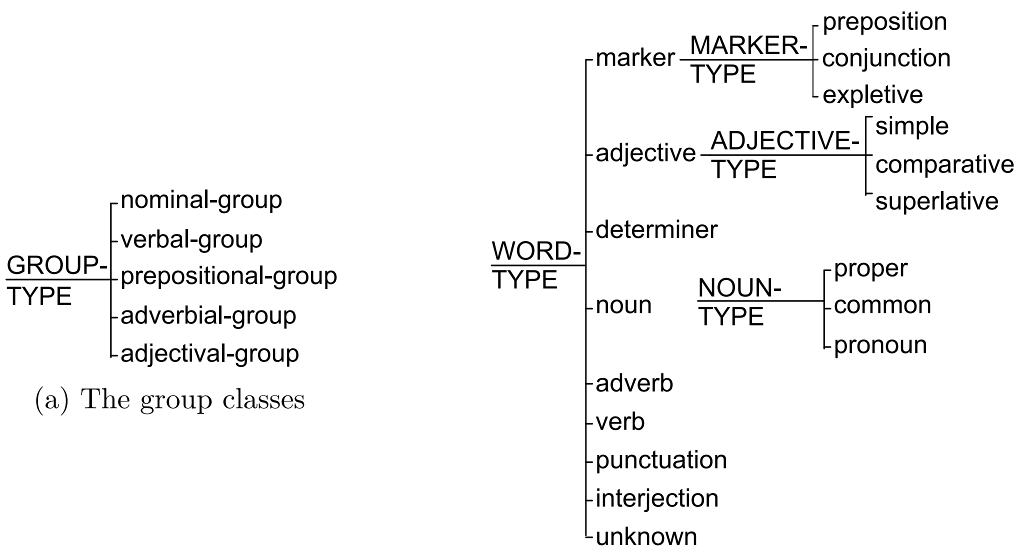
2.4.2 The unit classes

Fawcett drops the concept of rank system (discussed in Section 2.4.1) and through a bottom-up approach redefining the class as a "class of unit" as in 2.3.1.

He adopts Saussurean perspective on language which states that semantics and syntax are strongly intertwined with each other so major semantic classes of entities correspond to the major syntactic units. This lead Fawcett to take a semantic basis for classifying syntactic units into: clause, nominal group, prepositional group, quality group and quantity group (Fawcett 2000: p. 193–194) along with a set of minor classes such as genitive and proper name clusters.

The problem with this approach is that these classes are beyond the syntactic variations of the grammar and blend into lexical semantics which makes it difficult to apply to parsing, at least nowadays with current state of word classification.

In the current project I turn to Sydney classification of syntactic units that is close in line with traditional syntactic classification (Quirk et al. 1985). I adopt the clause as a unit plus the four group classes of ~~Sidney~~ grammar depicted in Figure 2.4.



(a) The group classes

(b) The word classes

Fig. 2.4 The group and word classes

2.4.3 The structure

The *unit* and *structure* are two out of the four fundamental categories in the systemic theories of grammar. Sydney and Cardiff theories vary in their perspectives on *unit* and *structure* influencing how units are defined and identified.

For Halliday, the *structure* (Definition 2.2.5) characterises each unit as a carrier of a pattern of a particular order of *elements*. The order is not necessarily linear realisation sequence but a theoretical relation of relative or absolute placement. This perspective has been ~~proved~~ useful in generation ~~exercises~~ where unit placement ~~evolves~~ ~~in~~ the realisation process.

The Cardiff School take a bottom up approach and defines class in terms of its internal structure describing a relative or absolute order of elements. This sort of syntagmatic account is precisely what is deemed useful in parsing and is the one adopted in this thesis. It is well established algorithmically how to recognise classes and construct them bottom up. So in our case easier to let the unit class emerge from recognition of constituent part-of-speech (word classes) and dependency relations between words or sequence of lower unit classes. In other words the unit class is defined by the unit structure and not by it's function in the parent unit, as Sydney school predicates, and this is precisely the reason why creation of constituency structure is computationally accessible.

2.4.4 Syntactic and semantic Heads

In SFG ~~the~~ heads may be semantic or syntactic. In most cases they coincide but there are exceptions when they differ or even ~~miss,~~ This is especially an important topic in the discussions of the nominal group structure on which Halliday & Matthiessen (2013) offer a thorough examination but Fawcett (2000) offers us a more generic perspective on this issue.

Consider the example of nominal group "a cup of tea" analysed in three different ways in the Table 2.5. The Sydney Grammar offers two analyses in which the semantic and the syntactic heads differ. In the *experiential* analysis the head is "tea" which functions as *Thing*, while in the *interpersonal* analysis the head is "cup" which functions as *Head*.

Cardiff Grammar does not make the Head/Thing distinction because the functional elements are already established based on semantic criteria. discussed in subsection 2.5.3. Nevertheless the logical analysis of SG resonates closely with the traditional "semantically blinded" grammars because it always provides a syntactic Head even if differs from the "pivotal element" of the group.

		a	**cup**	**of**	**tea**
Sidney Grammar	experiential	*Numerative*			*Thing*
	interpersonal	*Modifier*	*Head*	*Qualifier*	
Cardiff Grammar		*Quantifying Determiner*		*Selector*	*Head*

Table 2.5 Example of ~~dispersed~~ semantic and syntactic heads

Fawcett argues that none of the constituting elements of the unit is ~~mandatory~~ realised even the so called *"pivotal element"* which is the group defining element. The logical structure heads are always realised and correspond dependency relations established in the DG. Depending on the unit class logical structure heads are conflated with specific functions, for instance in nominal group the Head is usually conflated with the Thing, in quality group with the Apex, in quantity group with the Amount, in clause with the Main Verb and so on. But in language it is not unusual to have nominal groups with the Thing missing or elliptic clauses with missing the Main verb so no rigid correspondence can be established between the Head, unit class and the corresponding pivotal element of the group. So because the unit class depends on its internal structure ~~leading~~ to a circular interdependency between the unit class and the unit structure. To solve this issue Fawcett argues for ~~bototm~~-up approach where head-modifier relations are identified between lexical items and then between units (i.e. groups and clauses) serving as cues to identify elements of the higher unit and therefore it's class. Usually the class membership of head is raised to the unit class although sometimes the presence or absence of certain elements (during the reconstruction process) may alter the unit class to a different from the logical head.

(4) The old shall pass first.

Consider the nominal group "The old" which is the subject in example 4. The head of the nominal group is the adjective "old" and not a noun as ~~it~~ would normally be expected. The noun modified by the adjective "old" is left covert and it ~~shall~~ be recoverable from the context. We can insert a generic noun "one" to form a canonical noun group:"the old one". In such cases when the head noun is missing, the logical head is conflated with other element in this case the Epithet. The group class is not raised from the word class to quality group but is identified by internal structure of the whole group and in this case the presence of determiner signals a nominal class. I point out through this example that the class of the head is not always ~~is~~ raised to establish the group class but the whole underlying structure determines the group class.

2.4.5 Systems and systemic networks

For~~e~~ example consider polarity system represented in figure 2.5. It contains two choices either positive or negative. And when one says it is positive one means not negative which is obvious and self evident how the two choices are mutually exclusive.

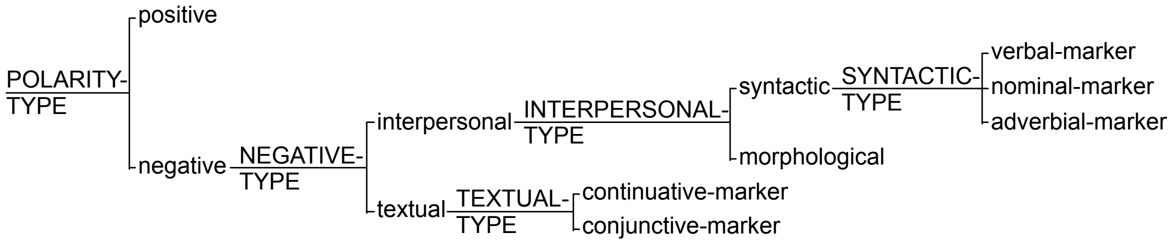


Fig. 2.5 System network of POLARITY

In language it may be often the case that a choice in a system may lead to re-entering the same system again to make another choice again forming this way recursive systems. Alternatively, we can say that a system allows multiple choices for the same unit. These perspectives are the sides of the same coin where the recursion perspective is useful for natural language generation while multiple choice perspective suits better the parsing task and next I explain why by continuing the discussion on polarity system.

The negative polarity in English clauses can be realised in several ways: via a noun group with intrinsic negative polarity feature like *"nobody"* (5), negation particle of the verb *"n't"* (6) or adverb with intrinsic negative polarity (7).

(5) Nobody with any sense is going.

(6) I don't have to mow my lawn.

(7) Never expect her to come back.

Consider now, the cases of double negation from the example 8 where two kinds of negations are realised in the same clause: the negation by verb particle *"n't"* and the pronoun *"nobody"*.

(8) Nobody with any sense isn't going.

(9) I don't have nobody to mow my lawn.

The systems can be recursive and thus choices are not always mutually exclusive. Even though the system network clearly distinguishes one type of negation from another multiple negations can still occurring simultaneously. Note that this more delicate distinctions in kind of negation, still is a negation and for any of them it is impossible to co-occur with positive polarity. The issue here is not semantic about whether the clause is positive or negative but what kinds of grammatical choices can be identified within the clause. The problem of whether the double negation shall be interpreted as positive is not necessarily as relevant as the task of identifying the two instances of negation.

Halliday states that the speaker makes only one choice from a system. If this rule is interpreted as two choices from the same system at a time being impossible then it clearly does not cover the recursive systems and needs weakening to accommodate border cases. I propose relaxing the constraint of *mutual exclusivity* to *disjunction*. Correspondingly, two types of systemic networks emerge differing by *the relation among choices*: the original Hallidayan XOR systems (such as POLARITY TYPE in figure 2.5) and the OR systems for accommodating cases of multiple feature selections (such as SYSTEMIC TYPE in the same figure).

As system is expanded in delicacy to form~~s~~ a systemic network of choices. Choice of a feature in one system becomes the entry condition for choices in more delicate systems below. I turn now to discuss the relationship types of relationships forming entry conditions to more delicate systems. For instance, an increase in delicacy can be seen as a taxonomic "is a" relationship between features of higher systems and lower systems like in the case of POLARITY TYPE and NEGATIVE TYPE in figure 2.5.

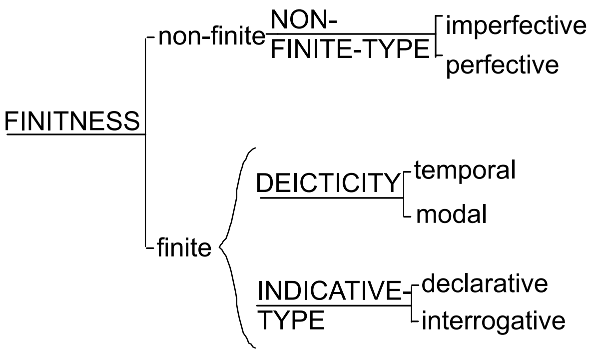


Fig. 2.6 A fraction of the finiteness system where increase of delicacy is not "is a" relation

The activation relation among systems in the cline of delicacy is not always taxonomic. Another relation is "enables selection of" without ~~a~~ any sub-categorisation implied. ~~For example~~ see FINITENESS system in figure 2.6 where in case that the finite option is selected then what this choice enables ~~are~~ not subtypes of finite but merely other ~~options~~ that become available i.e. DEIXIS and INDICATIVE TYPE. The latter is there because selection of finite implies also selection of indicative feature in a FINITNES's sibling system MOOD-TYPE comprised of options indicative vs. imperative.

In this subsection I defined the system and systemic network, presented two system types by the relationship between their choices and distinguished two kinds of activation relations between systems on the cline of delicacy.

2.4.6 coordination as unit complexing

In SG unit complexes fill an important part of the grammar along with the *taxis relations* which express the interdependency relations in unit complexes. *Parataxis* relations bind units of equal status while the *Hypotaxis* ones bind the dominant and the dependant units. Fawcett bypasses the taxis relations replacing them with coordination and embedding (Fawcett 2000: p. 271) seemingly an oversimplified approach leading to abandonment of unit complexing as well. While embedding accounts for the depth and complexity of syntax his approach to coordination is problematic. I ~~further~~ ~~discuss~~

~~and argue~~ for utility of unit-complexes for ~~the~~ coordination but this idea can be further extended to other phenomena which involve fixed idiomatic structures such as comparatives or conditionals.

Coordination is challenge not only for SFL but for other linguistic theories as well. CG treats this phenomena as two or more units filling or expounding the same element. For example, in table 2.6 "his shirt" and "~~and~~ his jeans" are two nominal groups that are siblings and both of them fill the same complement.

Ike	*washed*	*his*	*shirt*	*and*	*his*	*jeans*
Subject	Main Verb	Complement				
		Nominal Group		Nominal Group		
		Deictic Determiner	Head	&	Deictic Determiner	Head

Table 2.6 Coordination analysis in Cardiff Grammar

In SG the coordination is analised as a *complex unit* held together through paratactic relations ensuring that only one unit fills an element of the parent constituent in our example the complement of the clause. ~~The table~~ 2.7 illustrates an example analysis involving the complex unit approach.

Ike	*washed*	*his*	*shirt*	*and*	*his*	*jeans*
Subject	Predicate/Finite	Complement				
		nominal group complex				
		1		+2		
		Deictic	Thing	&	Deictic	Thing

Table 2.7 Coordination analysis in Sydney Grammar

~~The opinions~~ are divided by whether to invite the notion of a complex unit to handle coordination or not. If we dismiss the unit complex then an element could be filled by more than one units and if we adopt it then the complexing relations need to be accounted along with unit class and ~~what is~~ its structure.

I ~~would~~ argue for adoption of such unit type for two reasons. First, only units are accounted for structure while the elements can only be filled by an unit. Allowing multiple units to fill an element requires accounting at least for order if not also for the relation between the units. The structure as it is described in theories of grammar by Halliday (Halliday 2002) and Fawcett (Fawcett 2000) is defined in the unit and not the element. A unit has a specific possible structure in terms of places of elements however if an element is filled by two units simultaneously it constitutes a violation of the above principle as the order of those units is not accounted for but it matters.

(10) (Both my wife and her friend) arrived late.

(11) * (And her friend both my wife) arrived late.

(12) I want the front wall (either in blue or in green).

(13) * I want the front wall (or in green either in blue).

If the order would not have mattered then we could say that the conjunctions from ~~the~~ example 10 can be reformulated into 11 and the one from 12 into 13. But such reformulations are grammatically incorrect. Obviously the places do matter and they need to be described in the unit structure.

Secondly, the lexical items that signal the conjunction are not a part of the conjuncted units. This is contrary to what is being described in Cardiff and Sydney grammars. Fawcett present the Linker elements (&) which are filled by conjunctions as parts of virtually any unit class placed in the first position of the unit. Halliday omits to discuss in IFG (Halliday & Matthiessen 2013) the place of Linkers but implicitly proposes the same as Fawcett through his examples of paratactic relations at various rank levels where the lexical items signalling conjunction are included in the units they precede.

For example in the "or in green" the presence of "or" signalls the presence at least of one more unit of the same nature and does not contribute to the meaning of the prepositional group but to the meaning outside the group requiring presence of a sibling. Even more, the lack of a sibling most of the time would constitute an ungrammatical formulation. I say sometimes because it is perfectly acceptable to start a clause/sentence with a conjunction most often "but". But even in those cases it still invites the presence of a sibling clause/sentence preceding the current one to be resolved at the discourse level.

So conjunctions and pre-conjunctions shall not be placed as elements of the conjuncted units because they do not contribute to their meaning.

Adopting the unit complex and in particular coordination unit requires two clarifications (1) does the unit complex carry a syntactic class, and if so according to which criteria is it established? (2) Does it have any intrinsic features or not?

Zhang states in her thesis that ~~the~~ coordinating constructions do not have any categorial features thus there is no need to provide a new unit type. Instead the categorial properties of the conjuncts are transferred upwards (Zhang 2010). For example if two nominal groups are conjuncted then the complex receives the nominal class. This principle holds for most of the cases however there are rare cases when the units are of different classes. Consider the example 14 where the conjuncts are a nominal group “last Monday” and a prepositional group “during the previous weekend”.

(14) I lost it (either last Monday or during the previous weekend).

In this case there are two unit types that can be raised and it is not clear how to resolve this case. Options are to leave the class unspecified, transfer the class of the first unit upwards, or semantically resolve the class as both represent temporal circumstances even if they are realised through two different syntactic categories. Another option is to leave the class generic and assign the conjunction unit the class of *“coordination complex”* without sub-classifying it according to the constituent units below, i.e. without upward unit class transfer.

I address the second question regarding the intrinsic features of the complex unit. The coordination complex can have categorial features which none of the constituting units has. In ~~the~~ example 15 the conjunction of two singular noun groups requires plural agreement with the verb. Even though semantic interpretation that only one item is selected at a time, syntactically both items are listed in the clause and attempting third person singular verb forms ~~like~~ in 16 is grammatically incorrect.

(15) A pencil or a pen **are** equally good as a smart-phone.

(16) * A pencil or a pen **is** equally good as a smart-phone.

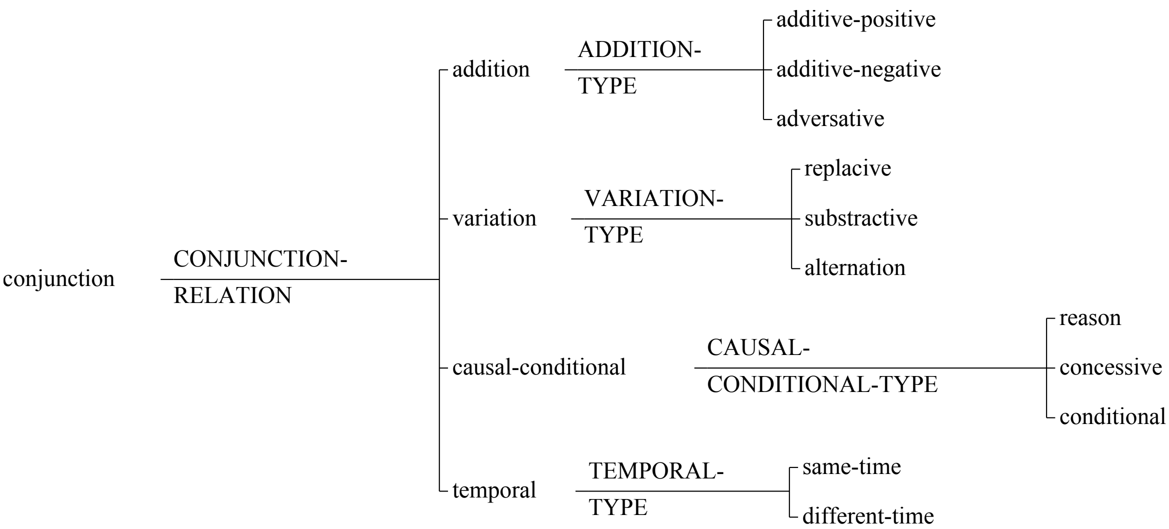


Fig. 2.7 Systemic network of coordination types

In the case of nominal group conjunction we can see that the plural feature emerges even if each individual unit is singular. For other unit classes it is not so obvious whether there are any linguistic features that emerge ~~to~~ the conjunction level. The meaning variation is ~~rather~~ semantic as for example conjunction of two verbs or clauses might mean very different things ~~like~~ consecutive actions, concomitant actions or presence of two states at the same time and so on. This brings us to another feature of the coordination complex - the type of the relationship it constructs. The lexical choice of head conjunct is the indicator of relationship among the conjuncts. Either *and*, *or*, *but*, *yet*, *for*, *nor* or *so* they express different meanings which are well representable as a relationship types systemic network in figure 2.7.

Adopting the unit complexing enables various kinds of constructions and coordination is only one of them. Below we discuss taxis relations and their role in unit complexing.

2.5 Critical discussion of the grammatical units

Now that the important theoretical details have been covered I would like focus on the grammars of the two schools. They have common parts and also differ in large parts on their paradigmatic and syntagmatic descriptions. This section discusses the main units considered in each of the grammars. ~~Like~~ in the previous section I argue on pragmatic grounds for adoption of unit structures from ~~either~~ grammar. This may seem like an inconsistency ~~add~~ below I ~~try to convince you of the opposite below~~. The argument runs along the line that some unit structures are closer to the syntactic analysis and thus are easier to detect and parse and the other ones may be a level of abstraction higher falling more on ~~the~~ semantic grounds thus becoming more difficult to capture in structural variance ~~and~~ requiring lexico-semantic resources.

For the reasons of limited space I skipped introducing the Sydney and Cardiff grammars and in turn assume that the reader is familiar with the details fo both of them. And for a general overview of the unit structure in each of the grammars ~~please~~ refer to Appendix 9.6. ~~Nevertheless, as it is a parallel contrastive discussion, even if the reader is familiar with one grammar only I hope it becomes clear how does certain phenomena are dealt with in the other one.~~

2.5.1 The verbal group and clause division

In SG the verbal group is described as an expansion of a verb just like the nominal group is the expansion of the noun(Halliday & Matthiessen 2013: p.396). There are certainly words that are closely related and syntactically dependent on the verb all together forming a unit that functions as a whole. For example the auxiliary verbs, adverbs or the negation particles are words that are directly linked to a lexical verb. The verb group functions as Finite + Predicator elements of the clause in mood structure and as Process in Transitivity structure.

In CG the verb group is dissolved moving the Main Verb as the pivotal element of the Clause unit. All the elements that form the clause structure and those that form the verb group structure are brought up together to the same level as elements of a clause. The clause structure in CG comprises elements with clause related functions(like Subject, Adjunct, Complement etc.) and other elements with Main Verb related functions(Auxiliary, Negation particle, Finite operator etc.).

Regard~~ing~~ from the Hallidayan rank scale perspective, merging the elements of the verb group into clause structure is not permitted because the units are ~~of~~ different rank ~~scales.~~ However it is not a problem for the relaxed rank scale version presented in subsection 2.4.1. The reason for adopting such an approach is best illustrated via complex verb groups with more than one non-auxiliary verb such as in examples 17–19.

~~Next I address~~ the impact of this merger on (a) the clause structure (b) the clause boundaries and (c) semantic role distribution within the clause.

(17) (The commission **started to investigate** two cases of overfishing in Norway.)

(18) (The commission **started** (**to investigate** two cases of overfishing in Norway.))

(19) (The commission **started** (**to finish** (**investigating** two cases of overfishing in Norway.)))

In SG "started to investigate" (example 17) is considered a single predicate of investigation which has specified the aspect of event incipiency despite the fact that there are two lexical verbs within the same verbal group. The "starting" doesn't constitute any kind of process in semantic terms but rather specifies aspectual information about the investigation process. The boundaries of the clause governed by this predicate stretch to entire sentence.

Semantically it is a sound approach because despite the presence of two lexical verbs there is only one event. However allowing such compositions leads to unwanted syntactic analysis for multiple lexical verb cases ~~like~~ in example 19.To solve this kind of problem~~s~~ Fawcett dismisses ~~the~~ verb groups and merges their elements into clause structure. He proposes the syntactically elegant principle of *"one main verb per clause"* (Fawcett 2008). Apply this principle to the same sentence yields a structure of two clauses illustrated in example 18 where the main clause is governed by the verb "to start" and the embedded one by the verb "to investigate". Note the conflict between "one main verb per clause" with Halliday's principle that only whole units form the constituency of others (the (c) principle of rank scale described in subsection 2.4.1). So allowing incomplete groups into the constituency structure would breach entire idea of unit based constituency.

Semantically the clause in SFL is a description of an event or situation as a figure with a process, participants and eventually circumstances where the process is realised through a lexical verb. Looking back to our examples, does the verb "to start" really describes a process or merely an aspect of it? Halliday treats such verbs as aspectual and when co-occurring with other lexical verbs are considered to form a single predicate. Accommodating Fawcett's stance, mentioned above and contradicting Halliday's approach, requires weakening the semantic requirement and allowing aspectual verbs to form clauses that contribute *aspectually or modally* to the embedded ones. I mention also the modal contribution because some verbs like *want, wish, hope*, etc. behave syntactically ~~exactly~~ like the aspectual ones. Moreover, Fawcett introduces into CG Transitivity network "influential" process type including all categories of meanings that semantically function as process modifiers: tentative, failing, starting, ending etc.

Fawcett's "one main verb per clause" principle changes the way clauses are partitioned, leads to abolition of the verbal group and introduces the "influential" process type.

2.5.2 The Clause

It is commonly agreed in linguistic communities that the unit of clause is one of the core elements in human language. It is consider the syntactic unit that expresses semantic units of a situation referring to a potentially rich array of meanings. The clause structure has been studied for long time and the main clause constituents are roughly the same in SFL as the ones in the traditional grammar (Quirk et al. 1985), transformational grammar (Chomsky 1957) an indirectly in dependency grammar (Hudson 2010).

In current work I adopt the CG Clause structure with the *Main Verb* as pivotal element. Though there is no element that is obligatorily realised in English ~~language~~, the Main Verb is realised with a reliably high degree of probability. Exceptions are the minor clauses (exclamations, calls, greetings and alarms) that occur in conversational contexts and elliptical clauses Halliday & Matthiessen (2013) such as the one in example 20 which are not currently covered.

(20) They were in the bar, *Dave in the restroom and Sarah by the bar.*

As mentioned before ?? the elements of a structure are defined in terms of their function contributing to the formation of the whole unit. The elements of an English clause are *Subject*, *Finite*, *Main Verb*(a part of the Predicator), up to two *Complements* and a various number of *Adjuncts*. All the elements of the assumed verbal group are part of the clause as well such as Auxiliary Verbs, Main Verb Extensions, Negators etc. (see Appendix 9.6 for a complete list).

2.5.3 The Nominal Group

The nominal group expresses things, classes of things or a selection of instances in that class. In ~~the~~ table 2.8 is presented ~~an example analysis~~ of the nominal group proposed in Sydney grammar (Halliday & Matthiessen 2013: pp. 364–369).

those	*two*	*old*	*electric*	*trains*	*from Luxembourg*
pre-modifier				head	post-modifier
Deictic	Numerative	Epithet	Classifier	Thing	Qualifyier
determiner	numeral	adjective	adjective	noun	prepositional phrase

Table 2.8 ~~The~~ example of a nominal group in Sydeny Grammar

In SG it is constituted by a head nominal item modified by descriptors or selectors such as: *Deictic*, *Numerative*, *Epithet*, *Classifier*, *Thing* and *Qualifier*. Each element has a fairly stable correspondence to the word classes, expected to be expounded by lexical items. ~~The~~ table 2.9 presents the mappings between the functions and the word classes.

Experiential function in noun group	**class (of word or unit)**
Deictic	determiner, predeterminer, pronoun
Epithet	adjective
Numerative	numeral(ordinal,cardinal)
Classifier	adjective, noun
Thing	noun
Qualifier	prepositional phrase, clause

Table 2.9 The mapping of noun group elements to classes

The elements in CG differ from those of SG. ~~The~~ table 2.10 exemplifies a noun group analysed with CG covering all the possible elements. ~~The~~ table 2.11 provides a legend for CG acronyms along with mappings to unit and word classes that can fill each element.

or	*a photo*	*of*	*part*	*of*	*one*	*of*	*the best*	*of*	*the*	*fine*	*new*	*taxis*	*in Kew*	*,*
pre-modifiers												head	post-modifiers	
&	rd	v	pd	v	qd	v	sd or od	v	dd	m	m	h	q	e

Table 2.10 The example of a nominal group in Cardiff Grammar

symbol	**function meaning**	**class (of word or unit)**
rd	representational determiner	noun, noun group
v	selector "of"	preposition
pd	partitive determiner	noun, noun group
fd	fractional determiner	noun, noun group, quantity group
qd	quantifying determiner	noun, noun group, quantity group
sd	superlative determiner	noun, noun group, quality group, quantity group
od	ordinative determiner	noun, noun group, quality group
td	tipic determiner	noun, noun group
dd	deictic determiner	determiner, pronoun, genitive cluster
m	modifier	adjective, noun, quality group, genitive cluster
h	head	noun, genitive cluster
q	qualifier	prepositional phrase, clause
&	linker	conjunction
e	ender	punctuation mark

Table 2.11 The mapping of noun group elements to classes in Cardiff grammar

The elements in CG ~~have~~ are based on semantic criteria supported by lexical and syntactic choices. Consequently some elements cannot be derived based on solely syntactic criteria requiring semantically motivated lexical resources. Semantically bound elements are predominantly determiners *Representational, Partitive, Fractional, Superlative, Typic Determiners* while the rest of the elements: *Head, Qualifier, Selector, Modifier and Deictic, Ordinative and Quantifying Determiners* can be determined solely on ~~the~~ syntactic criteria. The latter correspond fairly well to Sydney version of nominal group which is adopted in present work with the benefits of relaxed rank system replacing the sub-structures with embedded units and simplifying the syntactic structures.

Another simplification is renouncing to distinction between the Head and Thing (Halliday & Matthiessen 2013: p. 390–396) for the semantic ambiguity reasons as the determiners in CG. Thus if the logical Head of the nominal group is a noun then it is labelled as the Thing leaving the semantic discernment as a secondary process and out of the current scope. Otherwise, in cases of nominal groups without the Thing element, if the Head is pronoun(other than personal), numeral or adjective (mainly superlatives) then they function as Deictic, Numerative or Epithet. So I propose to parse the nominal groups in two steps: first determine the main constituting chunks and assign functions to the unambiguous ones and second perform a semantically driven evaluation for the less certain units.

In other cases the Thing is present but it is different from the Head. Consider example 21. In Sydney grammar ~~they're~~ treated as a nominal groups with qualifiers introduced by the *"of"* preposition. But these nominal groups are not really about the "cup", "some" or "another one" but rather about "tea", "youngsters" and "eruptions".

(21) (a cup) of (tea)

(22) (some) of (those youngsters)

(23) (another one) of (those periodic eruptions)

So then the syntactic Head still remains the first noun in the nominal group, but then by a semantic evaluation the Thing is shifted into the Qualifier introduced by "of" preposition. Cardiff Grammar weakens the assumption that every prepositional phrase acts as Qualifier in a nominal group and it the special case of the preposition "of". It is allowed to act not as the element introducing a prepositional phrase but as a end mark of a determiner-like selector. Thus making the former noun group a determiner in the latter one. This approach shifts the noun group head into the position of semantically based Thing and erases the discrepancy problem between them. Nonetheless this is not straight forward solution as it requires lexical-semantic informed decision. For example in 24 (Head/Thing marked in bold) the preposition "of" introduces Qualifiers.

(24) He was the **confidant** of the prime minister.

(25) It was the **clash** of two cultures.

The distinction between cases when the proposition "of" introduces a Qualifier or ends a Selector/Deictic requires a semantic evaluation answering the questions "what is the Thing that this nominal group is about?". While it is easy to just assume that the first noun in the nominal group is the head.

Therefore, I propose to parse the nominal groups in two steps: first determine the main constituting chunks and assign functions to the unambiguous ones and then in the second step to perform a semantically driven evaluation for the less certain units. This evaluation can be performed by further capturing the structure of nominal groups that act as Dyslectics through their lexico-syntactic realisation patterns.

2.5.4 The Adjectival and Adverbial Groups

Following the rationale of head-modifier ~~like~~ in the case of nominal groups, the adjectives and adverbs function as pivotal elements to form groups. The structure of adverbial and adjectival constructions is briefly covered in Sydney grammar in terms of head-modifier logical structures without an elaborated experiential structure like in the case of nominal groups. While the adverbial group is recognised as a distinct syntactic unit the adjectival group is treated as a special case of nominal group specifically as a sub-structure of Epithet or Classifier elements.

(26) He is *very lucky*.

For example "very lucky" in 26 is analysed as a short form of the nominal group "very lucky one" where "lucky" is the head of the nominal group with a missing Thing element "one". In this example "very" is not nominal modifier, it does not modify the missing nominal head but the adjective "lucky" so they constitute a head-modifier structure filling the Epithet element and as the rank scale system does not allow groups to fill elements of groups then it is described as a substructure of the nominal group.

In SG "The adverbial group has an adverb as Head which may or may not be accompanied by modifying elements"(Halliday & Matthiessen 2013: p. 419). The adverbial groups may fill modal and circumstantial adjunct elements in a clause corresponding to eight semantic classes of: time, place, four types of manner and two types of assessment. The adverbial pre-modifiers express polarity, comparison and intensification along with only one comparison post-modifier (Halliday & Matthiessen 2013: p.420-421).

A thorough systemic functional examination has been provided for the first time by Tucker Tucker (1997, 1998) materialised into a lexical-grammatical systematization of adjectives and the structure of Quality Group in CG. Fawcett uses the quality group structure for Adverbials as well as they follow similar grammatical behaviour. He avoids calling the group according to the word class but rather refers to the semantic meaning of what both groups express, i.e. the quality of things, situations or qualities themselves. The qualities of things have adjectives as their head while the qualities of situations an adverb.

In Cardiff Grammar, the head of the quality group is called *Apex* while the set of modifying elements: *Quality Group Deictic, Quality Group Quantifier, Emphasizing Temperer, Degree Temperer, Adjunctival Temperer, Scope* and *Finisher*. The quality group most frequently fills complements and adjuncts in clauses and fill modifiers and superlative determiners in nominal groups but there are also other cases found in the data.

Just like in the case of nominal group the adverbial and adjectival groups in Cardiff grammar are semantically motivated. To automatically identify elements of the quality group ~~requires~~ lexico-semantic resources.

Some adverbs are different from others at least because not all of them can be heads of the adverbial group like for example *very, much, less, pretty* also being able to act as adjectival modifiers whereas others cannot. So a naive attempt is to use a list of words to identify the Emphasizing and Degree Temperers.

Other elements of the quality group like the *Scoper* or *Finisher* are more difficult to identify and localize as part of the group only by syntactic cues and/or lists of words because of their inherent semantic nature. The problem is similar to detecting whether a prepositional phrase is filling a qualifier element in the preceding nominal group or ~~it~~ is filling a complement or adjunct in the clause. Not surprisingly the Scopers and Finishers are most of the time prepositional phrases.

Another issue is continuity. The question is whether a grammar shall allow at least at syntactic level ~~for~~ discontinuous constituents or not. And then if so how to detect all the parts of the group even if they do not stand in proximity ~~of~~ each other. For example, comparatives, a complex case of a quality group, could be realised in a continuous or discontinuous forms. Compare the analyses presented in 2.12 and 2.13. In the first case the comparative structure is a continuous quality group. In the second case the comparative is dissociated and analysed as separate adjuncts.

On one hand it is not a problem treating them as two adjuncts, because that is what they are from the syntactic point of view. However, semantically as Fawcett proposes, there is only one quality group with a discontinuous realisation whose Scope element is placed in a thematic position before the subject. For an automatic process

I	*am*	*much*	*smarter*	*today*	*than*	*yesterday*
Subject	Main Verb	Adjunct				
pronoun	verb	quality group				
		Emphasizing Temperer	Apex	Scope	Finisher	

Table 2.12 Comparative structure as one quality group adjunct

to identify a complex quality group is a difficult task. It needs to pick up ~~queues~~ like a comparative form of the adjective followed by the preposition “than” and then look for two terms being compared. Given some initial syntactic structure such patterns could be modelled and applied but only as a secondary semantically oriented process.

Today	*I*	*am*	*much*	*smarter*	*than*	*yesterday*
Adjunct	Subject	Main Verb	Adjunct			
adverb	pronoun	verb	quality group			
			Emphasizing Temperer	Apex	Finisher	

Table 2.13 Comparative structure split among two adjuncts

Since both the adverbial and adjectival groups have similar structures, it is syntactically feasible to automatically analyse them in terms of head-modifier structures in a first phase followed by a complementary process which assigns functional roles to the quality group components.

2.6 Discussion

This chapter ~~introduces~~ the fundamentals of systemic functional linguistics and ~~presents~~ an ~~adaptation~~ of Sydney and Cardiff theories of grammar ~~to~~ the task of parsing.

Because of bottom up approach to unit structure, rank scale relaxation and accommodation of embedding as a general principle Cardiff systemic functional theory is more suitable for parsing than Sydney one. Nonetheless the unit definitions in Cardiff grammar are deeply semantic in nature. Parsing with such units requires most of the time lexical-semantic informed decisions beyond merely syntactic variations. This is one of the reasons why the parsing attempts by O'Donoghue (1991) and others in COMMUNAL project were all based on a corpus.

As there was no corpus available and because the parsing approach is based on syntactic backbone none of the theories could be fully used as such. The second part of the chapter attempts to merge and adapt the grammars and theories of grammar to this parsing approach.

~~Next~~ chapter lays the theoretical foundations of Dependency Grammar and introduces the Stanford dependency parser used as a departing point in current parsing pipeline. Because there is a transformation step from dependency to systemic functional consistency structure, next chapter also covers a theoretical compatibility analysis and how such a transformation should in principle look like.