Systemic-Functional Linguistics and Computation: new directions, new challenges

John Bateman[1]   Daniel McDonald[2]   Tuomo Hiippala[3]
Daniel Couto-Vale[1]   Eugeniu Costetchi[1]

[1] Bremen University, Bremen
[2]Eberhard Karls Universität Tübingen
[3]Digital Geography Lab, University of Helsinki, Finland

Abstract

Systemic-functional linguistics (SFL) has a long history of interaction with computational linguistics, but no area stands still. In the 1980s and early 1990s, various computational systems for natural language generation informed by SFL achieved state-of-the-art performance. Subsequently, however, advances in statistically driven, rather than rule-based, computational linguistics eclipsed many of these earlier systems. In this contribution, we focus on several areas of contemporary computational applications of SFL to show how the field is now developing. We begin with a brief historical introduction to the context of this interaction, moving on with characterizations of the current state of the art and a consideration of the relationship between SFL and the statistical paradigm now central to computation and natural language processing.

1   Introduction

Computational linguistics (CL), natural language processing (NLP), and language technologies (LT) are closely intertwined fields that employ computational techniques for various tasks related to treatments of language. Key tasks include both the development of software tools (for information/document retrieval, document summarization, sentiment analysis, named entity extraction, corpus analysis and machine translation) and the development and refinement of linguistic theories via algorithmic models. The importance of algorithmic models became clear perhaps most significantly with Peters & Ritchie's (1973) formal proof that Chomsky's emerging transformational grammar was unlearnable in the form then under discussion. Results of this kind led to a new awareness that algorithmic properties could be deeply revealing of properties relevant for linguistic theorizing as well. More recently, work exploring the conditions under which language and language behavior can 'emerge' from situated interaction has also made considerable progress—progress that simply would not have been possible without the computational models necessary to conduct experimentation (Steels 2005).

Since the advent of computational approaches to language in the 1950s, computational models have in fact always been at the forefront of linguistic theorizing,

serving to push the degree of explicitness of such models further and revealing both gaps in treatments and new capabilities. Most grammatical theories have now received computational instantiation and there is extensive experience in designing algorithms for their processing, as well as substantial computationally accessible resources such as corpora, treebanks (i.e., grammatically annotated corpora), lexicons and more. The models employed usually attempt to ensure that their algorithmic properties are reasonable so that they do not suffer the same fate as early transformational grammar and fall foul of the unlearnability trap. It is this attention to computational properties that has allowed broad-coverage components that can be applied to real (large-scale) data to become the norm.

Given this, it should not then be surprising that systemic-functional linguistics (SFL) has a long history of interaction with computational linguistics (a detailed review of the historical engagement between the fields is given in O'Donnell & Bateman 2005). Bateman & O'Donnell (2015) trace the deep involvement of SFL's primary founder, Michael A.K. Halliday, in this process from its beginnings. Halliday participated in some of the earliest attempts to achieve automatic translation systems in the 1950s, bringing together linguistic theoretical considerations and practicalities of computational processing. He was also instrumental in shaping the linguistic foundations of some of the most well-known language-oriented systems to emerge in computational linguistics and Artificial Intelligence in the 1970s and 1980s. This included both Terry Winograd's SHRDLU (Winograd 1972), a landmark natural language dialogue system that demonstrated that natural dialogic interaction with computers was an achievable goal, and William C. Mann's Penman system for large-scale automatic natural language generation (Mann 1983; Mann & Matthiessen 1985).

From the 1980s and up until the mid-1990s, interaction between SFL and computation was consequently well established with other significant initiatives bringing the fields together. Robin Fawcett's COMMUNAL, for example, also approached automatic language generation, applying a different variant of SFL (Fawcett 1988), while Michael O'Donnell attempted to extend the capabilities on offer by developing automatic *analysis* components similarly based on SFL (O'Donnell 1994); many further systems are described in O'Donnell & Bateman (2005). There was, however, a marked difference in the relative successes and acceptance of these efforts. SFL, as a broadly functional theory of language focusing on language use as 'motivated choice', appeared well-suited to address natural language generation: here the abstract task is often characterized as precisely one of 'making the right choices' given a description of a language as a resource (cf. McDonald 1980). Most approaches to automatic text generation of that time, regardless of theoretical orientation, thus turned to questions concerned with finding the functional conditions under which particular (primarily) grammatical choices would be appropriate for the communicative goals being pursued. This characterization echoed directly that offered by Halliday concerning the main descriptive apparatus used within SFL, the system network (cf. Halliday 1996: 10).

The situation with analysis, i.e., the construction of computational components capable of moving from provided strings of words or sounds to more abstract, grammatical or semantic representations, was very different. Here there was very little success compared to the rapid growth of general purpose analysis systems in computational linguistics more broadly. The reasons for this asymmetry are themselves

of considerable theoretical note, revealing deeper issues in the formalizations of linguistic resources offered by the theory. The most significant of these are discussed in Bateman (2008b) and so will not be repeated here—essentially, however, questions of how to manage computational complexity of the kind mentioned above with respect to transformational grammar played a substantial role.

For the purposes of this chapter, what followed from the lack of success in the area of computational analysis using SFL specifications is more central. Whereas other approaches and systems began to achieve considerable success in automatic analysis, SFL's lack of success in this critical task led to it not being considered a viable approach for computation. This brought many consequences of its own, some of which were particularly important for the subsequent development of theory. For example, while SFL has always been oriented toward corpus-based work, pursuing such large-scale empirical research necessitates the availability or development of corresponding computational tools: one cannot examine large bodies of data by hand. Several other linguistic approaches have been more supportive of automated analysis of large-scale corpora and so it was natural that research—even research based on naturally occurring examples—would come to orient more to the kinds of theories for which analyzed corpora were, or could be made, available.

The success of these latter approaches had significant consequences for computational linguistics in general. As statistical methods, machine learning, automatic grammar construction and the like became ever more central to computational approaches to language throughout the 1990s, it became essential to create richly annotated datasets from which computational models could be derived. Machine learning, for example, works by taking a set of 'correct' examples and automatically deriving decision procedures capable of classifying previously unseen examples in the same way. To be effective, the quantity of 'correct' examples required to bootstrap the process can be quite large and SFL has simply not had resources of this kind—again largely due to a lack of automated analysis capabilities to get the entire process going. By the end of the 1990s, therefore, the position of SFL within computational linguistics had become relatively marginal.

Some legacy systems relying on SFL resources, such as the general natural language generation system KPML (Bateman 1997) descended from the Penman system, continued to be used and extended because of the considerable linguistic information they had come to include. For example, the development of the English grammar available with the KPML system stretches back to work by Matthiessen and Halliday in the 1980s in the Penman project (Mann & Matthiessen 1985); since then the grammar has come to include additions made by many further contributors giving rise to a grammar with very broad coverage—even by today's standards. The system as a whole thus came to occupy a particular niche among 'high-quality, high-effort' computational systems. More recently, less flexible but largely automatically produced generation systems relying on a variety of statistical methods have become common; an introductory overview of the field of natural language generation, its development over time, and current methods and challenges can be found in Bateman & Zock (2017).

Internally to SFL, the rapid developments in computational linguistics also began to have a significant, if largely indirect, impact. Corpus work, for example, frequently demands that bodies of data be 'marked up', or annotated, with particular categories that can subsequently be examined for meaningful patterns. This can

only sensibly be done using computational tools that support the annotation process and manage the corpus. Here Michael O'Donnell's UAM Corpus Tool (O'Donnell 2008) has been of enormous benefit to many researchers, operating both within and outside of SFL. This freely available program allows researchers to define their own classification systems and then supports application of those systems to bodies of text. In contrast to many such tools, UAM adopts the system network as its basic resource for defining classifications, thereby allowing deeply nested classifications of a kind particularly supportive of functional linguistic work. Pre-configured networks for commonly used areas of systemic-functional grammar are also provided.

Although the UAM Tool is intended to support manual annotation, i.e., annotation where the human researcher makes the choices of classification according to the options available, in its more recent instantiations it also provides access to some of the now standard computational components capable of producing structural analyses of unrestricted text, such as the freely available Stanford Parser (Manning et al. 2014). This now supports automatic structure analysis for several languages (including English, Arabic, French, German and Chinese). Because the Stanford Parser provides Phrase Structure and Universal Dependency (UD) parses, the relationship between its analyses and the categories of SFL is often far from straightforward. Currently, however, no comparable capabilities exist for systemic-functional grammars. This critical task is still unresolved from the perspective of SFL.

The general availability and wide-scale take-up of the UAM Tool demonstrates in addition how computational tools now form a normal part of the linguist's world and this is sure to increase as such tools gain even further in capabilities. Kay O'Halloran and team have, for example, produced a series of tools extending capabilities for corpus analysis both with respect to the depth of analysis, including semantics and discourse organizations, and to the breadth of analysis, moving into considerations of image, video and text-image combinations (O'Halloran 2014) as well. We return to this line of development below.

Ultimately, the current situation involving interactions between SFL and computation is complex. Whereas the lack of contact between computation and SFL by the early 2000s had led to a hiatus in new theoretical and practical engagements of SFL with computational techniques, the growing capabilities and sophistication of computational approaches to language have made that work increasingly relevant and difficult to ignore. As a consequence, there are now signs that a new revival in interaction is in progress. The availability of a far broader range of computational techniques, together with more accessible, robust and extensible infrastructures for developing and combining computational components, has made the development of new generations of computational SFL tools both possible and beneficial. This then forms the focus of the remainder of this article. We pick out several core areas in this newly emerging state of the art, describing current activities and identifying some key areas for future developments.

2   Parsing

We begin with the core task of providing SFL analysis capabilities, or 'parsing'. As mentioned above, the lack of such capabilities was one of the main reasons why interactions between SFL and computation faltered. Building a natural language parser can be seen as a task of creating an artificial text reader which understands

the meaning expressed in some text. The depth and the kind of text understanding varies according to the tasks addressed. Different levels of abstraction are required when enabling tasks such as summarizing documents, answering questions about them and their content, deriving new knowledge, interacting with a human user using natural language and so on. These and many other tasks are currently moving out of the field of artificial intelligence research and more and more into everyday life and practical applications.

Broad coverage natural language processing modules now exist for several levels of linguistic abstraction, ranging from the least abstract tasks of 'stemming' – i.e., removing inflectional information to reveal the basic lexical forms employed – and part of speech tagging, through intermediate tasks such as syntactic analyses, to highly complex tasks such as semantic analysis or argument extraction. But there are two caveats: in general, the higher the degree of abstraction, the less accurate the coverage becomes; moreover, the richer the linguistic description, the slower the parsing process. This is then particularly problematic for SFL because its grammars and other levels of description are rich and multi-layered in ways that differ from many other theoretical accounts.

More specifically, the descriptive power of a Systemic Functional Grammar (SFG) lies to a considerable extent in its separation of descriptive work across 'structure' (i.e., syntagmatic organizations) and 'choice' (i.e., paradigmatic organizations). This comes at the cost of high computational complexity, which still presents today the biggest challenge in parsing broad coverage texts with full SFGs. O'Donnell & Bateman (2005) discuss how each successive attempt to construct parsing components using SFL then necessarily led to the acceptance of limitations either in grammar size or in language coverage in order to proceed.

A parsing process for full SFGs needs then to derive both syntagmatic (e.g., constituency structure) and paradigmatic (i.e. selections from the system networks) descriptions. Providing a syntagmatic description is crucial for parsing as it is this organizational frame that serves as an anchor for structured paradigmatic details— that is, it is not sufficient to know that some feature has been selected; we also need to know precisely which grammatical unit that feature constitutes. Moreover, we need to be able to derive constraints on structure that are given by compatible feature selections and ruled out by incompatible feature selections. This latter task is a major source of computational complexity and, as Bateman (2008b) explains, brings with it significant theoretical implications for the construction of SFL theory as well as of computational systems. Today, it is common for parsers to rely on simpler syntactic trees (or other non-SFL grammars) as starting points for the parsing process. First, a syntagmatic organization, or 'structural backbone', would be defined, followed by an enrichment by paradigmatic selections. This technique was subsequently adopted as a beneficial heuristic for reducing complexity by most attempts to parse with SFGs (Kasper 1988; O'Donnell 2005; Costetchi 2013).

The first attempt to achieve larger-scale parsing capabilities for SFG was that of Robert Kasper (Kasper 1988). The structural backbone employed was provided by a context-free Phrase Structure Grammar (PSG), similar to Chomsky's use of a PSG to generate kernel sentences that would subsequently be subject to transformations (Chomsky 1957). In Kasper's case, each phrase-structure rule was given additional information for mapping the phrase structure onto a parallel systemic tree. After all possible systemic trees had been created, they were further enriched using infor-

mation from the Nigel Grammar (the SFG generation grammar developed within the Penman system—see Matthiessen 1985). This process was extremely slow and worked only on a limited size grammar because it involved 'multiplying out' all of the combinatorial possibilities inherent in the grammar. O'Donnell (1993) and Weerasinghe (1994) subsequently wrote parsers that attempted to build systemic functional constituency trees directly, avoiding the need to construct full phrase structure grammars. Again, however, the production of phrase structure rules for systemic clause structure was only feasible when limited grammatical possibilities were considered.

More recent approaches to syntactic parsing in computational linguistics more broadly have now advanced the field and offer far more powerful capabilities than the more experimental analysis components of the 1990s. It has therefore become logical to consider to what extent these may now better bootstrap the systemic parsing process as well. Costetchi (2013) consequently has designed and implemented a parser that still uses a syntactic backbone, but this time employing not a context-free phrase structure grammar but the Universal Dependency approach (UD—see Marneffe et al. 2014; Nivre 2015) used within the Stanford Dependency Parser (Marneffe et al. 2006; Socher et al. 2013). The approach remains broadly familiar: a structural backbone is derived using the broad coverage of the Stanford parser and this is then 'converted' into a form compatible with the systemic-functional syntagmatic organization.

Although the partially 'functional' nature of dependency relations provided by a UD parse can be used to support a more intuitive mapping to the functional elements defined in SFG, the conversion still raises challenges. UD is a single-layered grammar, oriented toward cross-linguistic validity and minimal redundancy, and as a consequence collapses features which in SFG would be assigned to different ranks (e.g., word class and roles within a clause, as in the 'nsubj' relation) and metafunctions (e.g., experiential Agents and interpersonal Subjects). The limited number of functional labels UD defines also means that UD descriptions distinguish far fewer cases than is commonplace within SFG. The 'nsubj' role provided by UD may correspond in an SFG to the Actor, Behaver, Sayer, Senser, Token, and Existent, among others. While the less delicate UD has obvious computational benefits (less data required for training, faster manual annotation, etc), it comes at the expense of descriptive breadth, especially considering the delicate orientation of many investigations carried out within SFL.

Costetchi's system approaches these difficulties as follows. First the parser transforms the dependency graph into a SFG constituency structure, specifically following the Mood structure, and afterwards enriches this with a series of features from Mood, Determination and Person system networks as described for SFG by Halliday & Matthiessen (2013). Next, the parser assigns process types and participant roles to constituents as defined in the Transitivity system network defined for Fawcett's (2008) variant of SFG. The assignment of these semantic features is not unique, however: the parser assigns all possible semantic configuration as opposed to the most probable one for the given clause, drawing on a database of process type structures (PTDB—see Neale 2002) and an auxiliary process that detects and creates placeholders for syntactically empty elements as described in Government and Binding theory (Haegeman 1991). Finally, for multiple clause sentences, the parser provides possible assignments of inter-clause tactic relations as described in Halli-

day & Matthiessen's (2013) system network for clause combinations. Transforming the dependency graph into a systemic Mood constituency tree and enriching it with features is performed by an implementation of a generic graph matching framework (Costetchi 2013). This allows descriptions of graph patterns with rich feature structures and embedded operations that are executed when their patterns matches a target. The same mechanisms based on graph patterns allow both enrichment with systemic features and creation of placeholders for syntactically null elements.

The work of Kasper (1988), O'Donnell (2005) and Costetchi (2013) highlight that approaches combining aspects of different grammars as well as cross-theoretic transformations are now practical and feasible options, especially when such computationally simpler grammars yield good performance and high levels of accuracy. This direction of research is then worth exploring further with other dependency grammars, such as the Link Grammar (Sleator & Temperley 1995), constituency grammars, such as HPSG (Collins 2003; Oepen et al. 2000) or XTAG (XTAG Research Group 2001), or Combinatorial Categorial Grammars (CCG—see Steedman 1993). All of these accounts are currently producing high-quality results with very broad coverage grammars. Other NLP tasks currently advancing in computational linguistics, such as semantic role labeling, temporal annotations, spatial annotations, named entity or concept identification and many others, might now also be incorporated employing similar mapping mechanisms (Costetchi 2013).

The task of providing full paradigmatic analyses remains challenging, however. This task can be formulated as a reasoning problem by treating existing systemic-functional grammars (such as the Nigel grammar) as a set of logical constraints constituting a combinatoric possibility space (of constituency structure, functions and descriptive features). The input of parses with other grammars can then be seen as 'factual' evidence to be considered when searching for solutions that resolve all the constraints given in the problem space. Although this abstract task still exhibits very high (computational) complexity, there is a long history in computer science considering solutions precisely to this problem (Kotthoff 2014). Particularly promising heuristics are offered by the problem reduction principle defined in computational complexity theory (Arora & Barak 2009) and the decomposition principle in probability theory (Grinstead & Snell 2012). Given a large high-dimensional distribution θ representing the domain knowledge, the task is to decompose it into a set of smaller lower-dimensional distributions $\{\theta_1, \theta_2...\theta_n\}$ from which the original distribution θ can be reconstructed with no errors. With such a decomposition one could draw any conclusions from $\{\theta_1, \theta_2 \ldots \theta_n\}$ that could be inferred from θ without actually reconstructing it.

This procedure already has implementations in terms of probabilistic graphical models (Airoldi 2007) using Bayesian Belief Networks and Markov Random Fields, which form the foundation for probabilistic logics such as Bayesian (Kersting & De Raedt 2007) and Markov Logic (Richardson & Domingos 2006; Domingos et al. 2010). They might therefore also offer good candidates for expressing system networks together with those networks' constraints on syntagmatic structure, especially given that highly efficient (polynomial and even log linear) learning and inference algorithms already exist for them (Guo & Hsu 2002).

Another direction worth exploring is building the syntactic backbone as native SFG constituency structures. For this task, an SFG corpus might be built either from scratch, such as the one in Fawcett's (1993) COMMUNAL project, or by trans-

formation of existing corpora (Honnibal 2004; Honnibal & Curran 2007). The latter option is more feasible and in line with the idea of cross-theoretic transformations across distinct grammatical accounts defended above. Such resources could then be employed in well established practice from computational linguistics as training datasets for machine learning algorithms, which we will return to briefly below.

Automated parsing with functional categories promises a number of important applications, both within and outside of research settings. Parsed data can help in determining the register dimensions of a text, assisting in document classification or analysis of diachronic change. These approaches can also help develop linguistic theory—in the case of SFL, automated frequency counting is perhaps the only feasible way of accomplishing the "grammarian's dream" (Hasan 1987) whereby grammatical distinctions and lexical alternatives become one unified resource. Such counting would enable accounts to systematically extend characterizations from system to instance respecting statistically-derived probabilities for given contexts.

Functional linguists approach corpora both from above (i.e., looking at collections of texts as assemblages of registers) and from below (i.e., by building profiles of lexicogrammatical frequencies). Teich et al. (2016), for instance, use register theory and selected elements of SFG to analyze a large, metadata-rich corpus of scientific writing. Automated tagging and manual annotation are used in tandem to extract frequency counts for various lexicogrammatical features. Statistical modeling is then used to model phylogenetic change, as well as disciplinary specialization. Findings show that disciplines differentiate themselves not only through experiential choices but through differing probabilities within tenor and mode as well. Teich et al. point out that tenor and mode are often neglected in tool and method development within NLP and SFL approaches could then help achieve more inclusive accounts.

Using similar methods, McDonald & Woodward-Kron (2016) investigate language change over the course of membership in an on-line support group for bipolar disorder. Over time, members' talk increasingly comes to align with a biomedical ideology: members' Mood choices shift in order to allow advice and foreground declaratives to provide hedged advice that foregrounds lay experience—experientially, users come to construe health problems as Possessions, rather than Identities. Zinn & McDonald (2015) apply a similar methodology in order to track shifting lexicogrammar and semantics of *risk* in print news journalism over the past 30 years. Focusing on experiential and group level features, they found that risk is increasingly nominal, negatively appraised, and construed as possible, rather than calculated.

In each of these contemporary projects, dedicated systemic parsing would radically increase the potential features chosen for analyses, and allow more precise division of meaning-making along metafunctional lines.

3 Dialogue systems: situated language use

Although the prospect of computational systems that converse with humans has always been upheld as one of the primary goals of artificial intelligence, progress has been relatively slow. The early system mentioned above, SHRDLU from Winograd (1972), was a landmark system that in fact proved a difficult act to follow. Over the past decade, an increasing number of computational systems with impressive dialogic capabilities have been produced, however, and the area is now coming back

into the mainstream: consumer electronics increasingly ship with computational 'assistants' of one kind or another, with each supporting at least spoken interaction between information systems and their human users.

When dialogue systems are constructed incorporating insights from linguistic theory, the interactional behavior that results offers a powerful source of additional knowledge and evaluation of the adequacy of the theories employed. It is generally immediately apparent when interaction does not run smoothly and so this can be considered in terms of whether the theories employed were adequate or not in the guidance they offered. This applies to all levels of abstraction within dialogue systems: for example, even when the speech recognition component is of insufficient quality to reliably resolve what was said acoustically, one might still expect the dialogue management component to respond more or less gracefully by politely inquiring again as to what was said, rather than simply failing with an internal system error.

The usual components of a computational dialogue system therefore span a considerable breadth of linguistic knowledge as well: ranging from spoken language recognition, parsing, semantic analysis, contextualization, recognition of speech acts, designing responses appropriate both to the context and to the addressee, producing grammatical forms for those responses and converting them into intelligible spoken output. Working on computational dialogue systems is then of considerable value for refining our linguistic theories in each of these areas *and* in combination. This applies equally to systemic-functional linguistics, and particularly to all of the varied components of the theory spread over levels of linguistic abstraction from phonetics to context of situation.

In general, developing a dialogue system can be seen as the task of creating an artificial 'persona', a non-organic being who will speak to humans. There are many reasons for developing such beings: some developers aim at alleviating human loneliness by building great listeners and companions while others focus on automating labour, work, and entertainment by implementing virtual call-center attendants, autonomous vehicles, vending machines, intelligent speakers, intelligent TVs, and intelligent home devices. To a greater or lesser extent, dialogues in all these contexts exhibit fundamental properties of 'situatedness'. Here, a host of well-known linguistic phenomena, such as deixis, i.e., referring to the speech situation, and recipient design, i.e., making sure that what is said is appropriate both to the context and to the state of knowledge of the addressee (cf. Fischer 2016), come to play central roles. Ensuring that aspects of the linguistic account are able both to access and to influence situation models appropriately to control these phenomena is then an important requirement.

These properties raise some particular theoretical challenges for systemic-functional theory that are currently unresolved: this concerns the entire area of establishing relations between the linguistic system and context. Although SFL has always placed considerable weight on the notion that language use, on the one hand, depends crucially on context and, on the other, plays a constitutive role in constructing such contexts, the mechanisms available within SFL for modeling this are still schematic. When building computational dialogue systems, however, precisely this property must be specified and implemented in detail. A major difficulty here is the highly 'dynamic' nature of the linkage between language and context. Each utterance is dependent on the context, while also changing that context for subsequent utter-

ances.

Approaches to dialogue systems outside of SFL most usually deal with this flexibility by combining the so-called 'information state' approach Traum & Larsson (2003), whereby semantic representations of the current 'question under discussion' and the states of knowledge of the respective interactional participants are used to trigger actions, and statistically derived probabilities concerned with which speech acts have occurred and which are most likely to follow. Within SFL-based systems, an early model of this interaction was set out by O'Donnell (1990), while Fawcett (1989) adopted more directly a procedural account based on flowcharts. More recent models interacting with systemic-functional descriptions also include levels of description concerned with transitions between dialogue states, modeled in a variety of ways (Teich et al. 1997; Shi et al. 2011), but these approaches still exhibit drawbacks in comparison to the non-SFL-based dialogue systems available. Non-SFL-based dialogue systems typically include far stronger formalizations of processes of reasoning with presuppositions, of the influence of the knowledge of addressees, of the communicative goals of the interactants, of the discourse history, of implications following from the semantics, and more besides. They can, therefore, by no means still be seen as impoverished or simplistic with respect to their purely linguistic competitors.

Even at the level of fine-grained lexicogrammatical choices, context-sensitivity of many linguistic phenomena must be catered for and theoretically characterized. Consider the case of a single proper name. Depending on the states of knowledge of the participants, uttering that name may be a simple mention, an invitation to participate, a direct address or call for action, and so on. The same participant may also be picked out by participant 'status'—i.e., according to whether the participant is a speaker (at some specific time), an addressee, or a third person participant or overhearer. Thus choices in the grammar need direct access to various organizational features of the situation and, moreover, those features change with each utterance and with time. Furthermore, depending on what precisely is being done with an utterance in a dialogue, the interpretation required may be quite different.

As a concrete example, consider the case of an autonomous wheelchair capable of spoken language interaction. Here, even uttered terms for objects, such as 'sofa' or 'TV', may function as the referenced object of a relative location playing the appropriate role in a command to move somewhere. In contrast, other uttered terms for objects, such as 'wheelchair', may serve as a way to direct a command at the intended addressee, as in 'Wheelchair, go to the kitchen'. Such variation demonstrates that anchoring linguistic analysis in the agency and affordances of things present in the situation is essential to determine the speech function of an utterance.

Situated dialogues therefore differ from monologues and from non-situated dialogues in substantial ways. On the one hand, a monologue such as an argumentative essay consists of a series of uncontested statements. As a result, each statement corresponds to a process that the author assumes took, is taking, or will take place. On the other hand, in a dialogue each statement must be acknowledged by the addressee(s) before interactants *agree* that something took, is taking, or will take place. Computational dialogue systems must then also incorporate explicit models of such 'grounding' as well in order to support natural interaction. Of course, interactants may also disagree about what can be accepted as having taken place,

which can lead to interruptions and further subdialogues at any point. In other words, while truth is proposed in monologues, in dialogues it is negotiated. In addition, information is not only offered by interactants, it is also demanded in the form of questions: in short, interactants exchange goods-and-services, and this must be explicitly modeled.

A computational model addressing the forms of flexibility described above at various levels has been developed by Couto-Vale (2017), building on SFL principles. The application domain for this system is that of the above mentioned 'intelligent wheelchair'. Such wheelchairs primarily play a role in interactive exchanges of goods-and-services. In particular, they offer a set of services, including going or taking someone to specified destinations and recharging themselves at an appropriate docking station. When interacting with such a wheelchair, a human normally gives the wheelchair a sequence of tasks for it to perform. These tasks are executions of the wheelchair's services. The context of a wheelchair offering services consequently motivates particular interpretations and descriptions of the utterances occurring. When humans and wheelchairs exchange services linguistically, the human takes the role of a service client and the wheelchair takes the role of a service provider. The human tells the wheelchair what to do and so is the source of the request; the wheelchair is the request's 'destination'.

This can then be captured in terms of functional configurations in the linguistic analyses that a dialogue system needs to perform when participating in natural dialogue in this scenario. Functional linguistic analyses appropriate for the context and constructed automatically on the basis of situation-specific resources are illustrated in Tables 1 and 2; in this scenario, the wheelchair goes by the name 'Rolland'. In Table 1, what is transferred from the service client (Agent) to the service provider (Medium) is a request, whereas in Table 2 what is described is a service of the service provider. Selecting the appropriate analysis for these grammatically very similar cases demands that a language analyzer knows of the respective capabilities of the represented participants: only the wheelchair is capable of moving the client, whereas only the client is capable of calling for services.

I	called	Rolland	to my bed
I	brought	Rolland	to my bed
Client	Process	Provider	Route
Agent	–	Medium	Update

Table 1: Description of request (by human)

Rolland	took	me	to my bed
Rolland	brought	me	to my bed
Provider	Process	Client	Route
Agent	–	Medium	Update

Table 2: Description of service (by human)

Following this line of discussion further: not every process represented in human-wheelchair interaction is a service, and so the system, and the underlying theoretical description, must be able to tease these distinct interpretations apart. For example, humans also use descriptions of their own future actions in order to communicate to the wheelchair that services need to be performed that enable users' actions to take place. For instance, let us assume a gait-impaired woman is interacting with an intelligent wheelchair and wishes to wash her hands. In this situation, only the woman can wash her hands; the wheelchair cannot. However, the woman can only wash her hands if she is at a given position in relation to a sink. Resolving the intention of an utterance such as 'I must wash my hands' is therefore complex and

again demands that the system can access just who can perform what actions in order to assign functional roles in the grammatical analysis appropriately.

Specifically, because the wheelchair offers services it needs to construct interpretations even of representations of human actions (such as 'I must wash my hands') in terms of possible actions that would support those actions. It is this that makes it possible for the wheelchair to respond in the present case: "ok, then I'll take you to the sink". In other words, it is the fact that interactants can perform some actions and cannot perform other actions, and that interactants offer services that enable each other's actions, that enables such clauses to be construed as commands for the wheelchair to do something. Such knowledge must then be accessible if situationally appropriate interpretation of utterances is going to be possible (Couto-Vale 2017).

Several practical approaches to implementing computational dialogue systems of this kind make simplifying assumptions that provide basic functionality but with little scope for extension to cover more natural or complex interactions of the kind illustrated here. For example, spotting keywords is a straightforward way of 'guessing' what an utterance might have meant. When the word 'kitchen' is spotted somewhere in an utterance, the wheelchair or other assistant may just assume that the intended service is to take the user to the kitchen. This strategy becomes increasingly unwieldy when extended to interpret recognition of a keyword such as 'wash' as an instruction to go somewhere. What happens in practice with such straightforward approaches is that multiple keywords such as 'wash' + 'hands', 'wash' + 'hair', 'wash' + 'dishes', and 'wash' + 'clothes' are necessary for guessing the command. These sets of keywords become the conditions of interpretation rules, which get progressively more complex as the number of potential clauses increase. For this reason, although such an implementation strategy works for simple scenarios, it does not result in reusable linguistic resources nor in resources that are easy to maintain as the domain of application expands.

Being able to describe features of utterances as systemic options with reference to the situation as suggested here and following SFL principles then promises an economical way of automating understanding and generation in general. It is by recognizing the features of the situation that enable and disable people to mean something in particular that we can produce resources for one situation and reuse them in another. Developing more complex dialogue systems building on fine-grained and dynamic linkages between language and situation is in fact not only of interest for applications; it is also a highly effective strategy for pushing theory development. When a computational system is brought to the level of explicitness and completeness that it can actually produce behavior (at any levels of linguistic abstraction), weaknesses or gaps in the theories implemented can become glaringly evident in a way that is simply not accessible when considering the theories 'on paper'. When the produced behavior does not meet expectations, this is a good indicator of places where theoretical frameworks may need refinement.

4   Multimodality

Another development gaining momentum in several areas of computational linguistics is the focus beyond language to include other modalities or forms of expression. With this move, the range of computational work relevant for non-computational theory building is also extended considerably—particularly for the area of multi-

modality. Multimodality is the study of how multiple modes of expression interact with each other in diverse communicative situations (Bateman et al. 2017). To draw an example: a multimodal approach to the dialogue systems discussed above might pay attention to gestures alongside spoken language, thus extending the scope of relevant work from natural language processing to gesture recognition, a subfield of human–computer interaction (Rautaray & Agrawal 2015). Alternatively, a multimodal perspective on page-based documents might in turn involve document analysis, a subfield of computer vision, to automatically infer the structure of the document and identify its contents, and apply appropriate processing depending on whether they consist of photographs, diagrams, other graphic elements or written language (Doermann & Tombre 2014) and to derive text-image relations (Bateman 2014).

While this illustrates how issues of multimodality and computation can often be intertwined, a discussion capturing their breadth is well outside of the scope of this chapter. Therefore, to limit and structure the discussion on computational methods in SFL-inspired work on multimodality, we begin with a body of work strongly rooted in SFL, namely that directed by Kay O'Halloran, before considering how recent advances in computational methods may benefit and inform future research on multimodality.

In work undertaken at the Multimodal Analysis Lab at the National University of Singapore between 2008 and 2013, O'Halloran and her team aimed to:

> "... develop and use interactive digital technologies for multimodal analysis of different media and to develop computational, mathematical, and visualization techniques for interpreting semantic patterns in the resulting multimodal data. The research program also aimed to develop automated computational techniques for analysis of large cultural data sets, and to develop digital technologies that promote a systematic approach to teaching and learning multimodal literacy and communication skills for the twenty-first century." (O'Halloran 2015: 390)

Much has been written elsewhere about the tools for supporting multimodal analysis (see e.g. O'Halloran et al. 2014b); in the following, therefore, we focus on the application of computational techniques to analyzing multimodal data.

Despite the computational emphasis, the underlying theoretical framework of O'Halloran's projects draws heavily on what O'Halloran (2008) conceptualizes as Systemic-Functional Multimodal Discourse Analysis (SF-MDA). Following the social semiotic and systemic-functional approaches to multimodality (e.g. Kress & van Leeuwen 2006; O'Toole 2011), SF-MDA considers language and image as resources for making-meaning, building on the rich theoretical framework provided by systemic-functional theory. This involves, for instance, applying the concepts of metafunctions and rank to visual images, so as to provide an integrative framework for describing multimodal data. With this kind of framework at hand, another question quickly emerges from the computational perspective: namely, how to move beyond hand-picked examples and bring the powerful theoretical apparatus of SFL to bear on large volumes of multimodal data. Indeed, this question is well-known and long-discussed in non-SFL computational corpus analysis.

Similar questions are posed in the emerging field of digital humanities, which studies how computational methods and techniques can help to answer research

questions raised in the humanities (see e.g. Berry 2012; Schreibman et al. 2016). Multimodality, however, has not been theorized extensively in the digital humanities, except in discussions of developing new ways of representing and disseminating research results (Svensson 2010). O'Halloran et al. (2014a: 565) consequently characterize their work as a further extension into what they define as "multimodal digital humanities", which involves collecting large volumes of linguistic and visual data, which are then aggregated for analysis and converted into interactive visualizations for exploration.

O'Halloran et al. (2014a) provide a good example of this approach at work, exploring the dynamics of urban life in Singapore. To study large volumes of data collected from various social media services, O'Halloran et al. (2014a) examined interpersonal meanings across two different semiotic resources by using computational techniques. For written language, O'Halloran et al. (2014a: 572) evaluated the sentiment of 98,733 Twitter messages sent from specific locations by calculating a lexicon-based emotion vector for each message, which captured the basic emotions of happiness, sadness, fear, anger, disgust and surprise. For photographs, they applied automatic face detection to 301,865 images retrieved from Instagram (O'Halloran et al. 2014a: 573). By aggregating this information into a grid defined over the map of Singapore, in which the locations are enriched with venue information combined from FourSquare and Wikipedia, O'Halloran et al. (2014a) investigated differences between residential and tourist areas.

In another study, Cao & O'Halloran (2015) explored differences in photo-shooting patterns between different groups of users on Flickr, focusing on differences in shot scale (close-up vs. long distance). Trained using texture patterns extracted by computer vision algorithms from the photographs, a machine learning algorithm—a Support Vector Machine (SVM)—learned to distinguish between shot scales with 91.3% accuracy. Cao & O'Halloran (2015) then applied this classifier to examine photographs taken by different groups. This revealed a strong correlation between the user's geographic location and shot scale: the users were more likely to take close-up shots while in their home country, while taking more distant shots abroad. Cao & O'Halloran (2015) suggest that this may result from individuals taking pictures of more mundane objects at home, while capturing sights and sceneries abroad, reflecting photographic practices associated with tourism.

Although the results discussed above show the potential of computational methods, their level of detail remains far from those commonly found in manual analyses of multimodality within SF-MDA and other approaches. O'Halloran et al. set out to bridge this gap in subsequent work, identifying the following challenges:

> "First, it is not possible to model and predict discourse patterns extrapolating from a limited number of detailed analyses. Second, the modeling of multimodal data using dimensionality reduction and clustering techniques results in visual patterns that require a human analyst to make sense of them, rather than delivering explicit, computable accounts of the semantic patterns which have been derived." (O'Halloran et al. 2016a: 10)

To this end, they propose a "multimodal mixed methods research framework" that uses qualitative methods to identify key semiotic resources in the collected data set, which are then targeted using quantitative methods, such as mining the data with

the help of machine learning algorithms. Finally, these analyses are synthesized into visualizations for exploring the results.

As the work of O'Halloran and her colleagues shows, rapidly developing fields of study such as computer vision, natural language processing and machine learning will undoubtedly make a significant contribution to the study of multimodality in the coming years. The work of Bateman et al. (2016) exemplifies emerging work in this area, combining manually and automatically generated annotation layers in a corpus describing the multimodality of film. Whereas various visual and aural events in film, such as shot boundaries and background music, are captured automatically by algorithms, filmic cohesion is described manually using the framework set out in (Tseng 2013).

Bateman et al. (2016) show that automatically generated annotation not only reduces the time and resources spent on compiling multimodal corpora, but also extends their scope by introducing layers of description which could be otherwise considered too demanding for manual annotation. Moreover, these benefits are not limited to the description of complex dynamic multimodal phenomena in film, but apply to page-based artifacts as well, whose annotation has proven equally time- and resource-intensive. To draw on an example, Hiippala (2016) presents a tool which uses open source computer vision, natural language processing and optical character recognition libraries to generate XML annotation from document images, which is designed to support the manual application of the annotation framework presented in Bateman (2008a).

Developments of this kind are now being driven forward by advances in several interrelated areas of computer science, including machine learning, computer vision and NLP. In particular, results in a specific subfield of machine learning known as 'deep learning' are now bringing about significant developments in all of the aforementioned fields. This subfield focuses on the design and use of artificial neural networks for a broad variety of tasks (LeCun et al. 2015). Artificial neural networks follow principles of operation inspired by the structure and activation of neurons in the human brain and have become increasingly popular in recent years due to increases in computing power and the volume of data available for training the networks.

In contrast to many 'traditional' machine learning algorithms, in which the features necessary for the task at hand—such as classifying an image based on its contents or finding an object in the image—are crafted manually by humans, neural networks learn these representations automatically by adjusting their parameters during training. The epithet 'deep', in turn, refers to how these parameters are organized into successive layers, which learn increasingly abstract representations of the data. Whereas the first layers may contain representations of changes in illumination or texture, the subsequent ones may construe combinations of these features into representations of particular objects. These developments are highly relevant to multimodal research, as exemplified by now common computer vision tasks such as image captioning and object detection, recognition and classification (see Bateman et al. 2017: 162–166).

As a concrete example, the work of Kembhavi et al. (2016) focuses on understanding the content and structure of diagrams and shows how far deep learning has pushed the joint processing of language and images. Kembhavi et al. (2016) train multiple neural networks to parse the diagrams for constituents and their rela-

tionships. This representation of the diagram is then fed to another neural network for 'diagram question answering', which involves predicting the correct answer to a multiple choice question. How diagrams make meanings has also been explored from an SFL-inspired, multimodal perspective by Guo (2004), but as in many other cases, these analyses have been limited to very few examples. In contrast, the parser presented in Kembhavi et al. (2016) could be used to radically increase the size of multimodal corpora for diagram description. In fact, their dataset, which includes 5,000 human-annotated diagrams, warrants a multimodal investigation in its own right.

Finally, it is useful to consider the potential contribution of SFL and multimodal research to what may be broadly described as the field of artificial intelligence. Although O'Halloran et al. (2016b) envisage that multimodal analyses conducted by experts could be re-used as training data for machine learning algorithms, state-of-the-art techniques such as deep learning are notoriously hungry for data. This data is needed both for training models for various tasks and to measure their performance. Typically, this data is crowd-sourced through services such as Amazon Mechanical Turk, which allows individuals to bid on and undertake small tasks, such as labeling objects or drawing their outlines in images. At this stage, given their experience in developing rich and systematic annotation schemes, researchers working on multimodality could develop annotation frameworks for multimodal data and investigate how best to instruct the non-expert annotators in their task, thus improving the quality of data.

5 Toward further integration

Better engagement with CL and NLP has a number of benefits for SFL. First, computational methods can facilitate automatable and reproducible work. The large amounts of time taken for manual annotation of data mean that many SFL projects face time and cost constraints. Heavily automated workflows, on the other hand, can be deployed on new data at little to no cost. This seems to be a practical path to Matthiessen's notion of language as "an assemblage of registers" (Matthiessen 2015b: 44): the same set of routines, automatically applied to corpora of different domains, could provide an insightful account of how Field, Tenor and Mode influence the probabilities for content-stratum phenomena.

At the same time, computational approaches make it possible to empirically test key components of an SFG. Automated text processing, for example, may be able to shed light on the oft-noted difficulty of process-type identification: if a model trained on large, well-annotated collections of process types cannot accurately predict process type labels in unseen data from a similar text-type, we may have an indication that our current understanding of experiential semantics is incomplete. Moreover, as Halliday discovered, the number of words needed to collect a quantitatively useful sample grows exponentially with the delicacy of the phenomenon of interest. While only 2000 main clauses are needed to create a profile of Mood or Process Type, hundreds of thousands of words (or a smaller, highly specific sample) may be needed in order to develop frequency profiles for lexical alternatives (Matthiessen 2015a). Realizing the statistical component of SFG at the grammatical pole of lexicogrammar is therefore dependent on computational methods.

Another important benefit of combining SFL and NLP is that the high-quality

data produced by human annotators with detailed training in SFG could be effectively repurposed as training data for machine learning algorithms. As noted above, a consistent, well-annotated dataset is the main prerequisite for the development of a high-quality statistical parser.

Finally, for NLP practitioners, SFL provides several well articulated hypotheses concerning the relationship between critical linguistic questions, such as the relationship between text and context, lexis and grammar, and words and meaning. In clearly differentiating between form, function, words and meaning, SFG may be able to avoid pitfalls that limit the utility of more popular computational grammars for functional-semantic tasks. It is certainly possible that current limitations in NLP are not the result of limitations in statistical methods, but in the grammars accepted by algorithms as input and output. Put another way, even perfectly accurate automatic annotation may have limited usefulness if what is being annotated does not correspond to meaningful distinctions within the grammar of a language. Goals that are still distant in NLP, such as semantic parsing and discourse-level annotation, could foreseeably stand to benefit from the relatively holistic account provided by SFL.

6   Conclusion

SFL is a complex and rich theoretical account in which several distinct representational resources are regularly used. Many of these are currently at the limit or beyond what can be modeled computationally. Computation has, however, made enormous strides over the past 10 years and many of the phenomena at the core of SFL theorizing at the outset—particularly, for example, reliance on data and corpora—are now coming within reach. In fact, such approaches to language are only achievable with computational methods and so it is both necessary and logical that interaction between SFL and computation should continue and, in some areas, intensify. In previous rounds of interaction the efforts of particular individuals have been central. For example, beginning with work on machine translation and then, later, natural language generation, the willingness of Michael Halliday to engage with emerging technologies played a crucial role (cf. Bateman & O'Donnell 2015). There is now considerably more need for researchers who are trained both in computation and SFL as both of these disciplines continue to evolve and grow.

References

Airoldi, Edoardo M. 2007. Getting Started in Probabilistic Graphical Models. *PLoS Computational Biology* 3(12). e252. doi:10.1371/journal.pcbi.0030252. < http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2134967/>.

Arora, Sanjeev & Boaz Barak. 2009. *Computational complexity: a modern approach.* Cambridge University Press.

Bateman, John & Michael Zock. 2017. Natural Language Generation. In R. Mitkov (ed.), *Oxford Handbook of Computational Linguistics*, Oxford: Oxford University Press 2nd edn.

Bateman, John A. 1997. Enabling technology for multilingual natural language generation: the KPML development environment. *Natural Language Engineering* 3(1). 15–55.

Bateman, John A. 2008a. *Multimodality and Genre: a foundation for the systematic analysis of multimodal documents.* Basingstoke: Palgrave Macmillan.

Bateman, John A. 2008b. Systemic functional linguistics and the notion of linguistic structure:

unanswered questions, new possibilities. In Jonathan J. Webster (ed.), *Meaning in context: strategies for implementing intelligent applications of language studies*, 24–58. London: Equinox.

Bateman, John A. 2014. *Text and Image: a critical introduction to the visual/verbal divide*. London and New York: Routledge.

Bateman, John A. & Michael O'Donnell. 2015. Computational Linguistics: the Halliday Connection. In Jonathan J. Webster (ed.), *The Bloomsbury Companion to M.A.K. Halliday*, 453–466. London and New York: Bloomsbury.

Bateman, John A., Chiao-I Tseng, Ognyan Seizov, Arne Jacobs, Andree Lüdtke, Marion G. Müller & Otthein Herzog. 2016. Towards next-generation visual archives: image, film and discourse. *Visual Studies* 31(2). 131–154. doi:http://dx.doi.org/10.1080/1472586X.2016.1173892.

Bateman, John A., Janina Wildfeuer & Tuomo Hiippala. 2017. *Multimodality – Foundations, Research and Analysis. A Problem-Oriented Introduction*. Berlin: Mouton de Gruyter.

Berry, David M. (ed.). 2012. *Understanding digital humanities*. Basingstoke: Palgrave Macmillan.

Cao, Yanpeng & Kay L. O'Halloran. 2015. Learning human photo shooting patterns from large-scale community photo collections. *Multimedia Tools and Applications* 74(24). 11499–11516. doi:10.1007/s11042-014-2247-0.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton & Co.

Collins, Michael. 2003. Head-driven statistical models for natural language parsing. *Computational linguistics* 29(4). 589–637.

Costetchi, Eugeniu. 2013. A method to generate simplified Systemic Functional Parses from Dependency Parses. In *Proceedings of depling2013 [forthcoming]*, Prague.

Couto-Vale, Daniel. 2017. *How to make a wheelchair understand spoken commands*. Bremen: Bremen University dissertation.

Doermann, David & Karl Tombre (eds.). 2014. *Handbook of document image processing and recognition*. London: Springer-Verlag.

Domingos, Pedro, Stanley Kok, Daniel Lowd & Hoifung Poon. 2010. Markov Logic. *Journal of computational biology a journal of computational molecular cell biology* 17(11). 1491–508. doi:10.1089/cmb.2010.0044. < http://www.ncbi.nlm.nih.gov/pubmed/21685052>.

Fawcett, Robin P. 1988. Language generation as choice in social interaction. In M. Zoch & G. Sabah (eds.), *Advances in natural language generation*, 27–49. London and New York: Pinter.

Fawcett, Robin P. 1989. Towards a systemic flowchart model for discourse analysis. In Robin P. Fawcett & David Young (eds.), *New developments in systemic linguistics: theory and application*, 116–143. London: Pinter.

Fawcett, Robin P. 1993. The architecture of the COMMUNAL project in NLG (and NLU). In *The Fourth European Workshop on Natural Language Generation*, Pisa.

Fawcett, Robin P. 2008. *Invitation to Systemic Functional Linguistics through the Cardiff Grammar*. London: Equinox Publishing Ltd.

Fischer, Kerstin. 2016. *Designing Speech for a Recipient: The roles of partner modeling, alignment and feedback in so-called 'simplified registers'*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Grinstead, Charles Miller & James Laurie Snell. 2012. *Introduction to probability*. American Mathematical Soc.

Guo, Haipeng & William Hsu. 2002. A survey of algorithms for real-time Bayesian network inference. In *Aaai/kdd/uai02 joint workshop on real-time decision support and diagnosis systems*, Edmonton, Canada.

Guo, Libo. 2004. Multimodality in a biology textbook. In Kay L. O'Halloran (ed.), *Multimodal discourse analysis: systemic functional perspectives* Open Linguistics Series, 196–219. London: Continuum.

Haegeman, Liliane. 1991. *Introduction to Government and Binding Theory*, vol. 2. London: Blackwell.

Halliday, Michael A. K. 1996. On grammar and grammatics. In Ruqaiya Hasan, Carmel Cloran & David Butt (eds.), *Functional descriptions – theory in practice* Current Issues in Linguistic Theory, 1–38. Amsterdam: Benjamins.

Halliday, Michael A.K. & Christian M.I.M. Matthiessen. 2013. *An Introduction to Functional Grammar (4th Edition)*. Oxford, New York: Routledge 4th edn.

Hasan, Ruqaiya. 1987. The grammarian's dream: lexis as most delicate grammar. In Michael A. K. Halliday & Robin P. Fawcett (eds.), *New Developments in Systemic Linguistics. Volume*

1, 184–211. London: Frances Pinter.

Hiippala, Tuomo. 2016. Semi-automated annotation of page-based documents within the Genre and Multimodality framework. In *Proceedings of the 10th sighum workshop on language technology for cultural heritage, social sciences, and humanities*, 84–89. Berlin, Germany: Association for Computational Linguistics.

Honnibal, Matthew. 2004. Converting the Penn Treebank to systemic functional grammar. In *Proceedings of the Australasian Language Technology Workshop (ALTW04)*, <http://www.alta.asn.au/events/altw2004/publication/04-27.pdf>.

Honnibal, Matthew & James R. Curran. 2007. Creating a Systemic Functional Grammar Corpus from the Penn Treebank. In *Proceedings of the ACL 2007 Workshop on Deep Linguistic Processing*, 89–96. Prague, Czech Republic: Association for Computational Linguistics. <http://www.aclweb.org/anthology/W/W07/W07-1212>.

Kasper, Robert. 1988. An Experimental Parser for Systemic Grammars. In *Proceedings of the 12th International Conference on Computational Linguistics*, Budapest.

Kembhavi, Aniruddha, Mike Salvato, Eric Kolve, Minjoon Seo, Hannaneh Hajishirzi & Ali Farhadi. 2016. A diagram is worth a dozen images. In *Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV*, 235–251. Cham: Springer.

Kersting, Kristian & Luc De Raedt. 2007. Bayesian Logic Programming: Theory and Tool. In Lise Getoor & Ben Taskar (eds.), *Introduction to statistical relational learning*, chap. 10, 291–321. MIT Press.

Kotthoff, Lars. 2014. Algorithm selection for combinatorial search problems: A survey. *AI Magazine* 35(3). 48–60.

Kress, Gunther & Theo van Leeuwen. 2006. *Reading Images: the grammar of visual design.* London and New York: Routledge 2nd edn.

LeCun, Yann, Yoshua Bengio & Geoffrey Hinton. 2015. Deep learning. *Nature* 521. 436–444. doi:10.1038/nature14539.

Mann, William C. 1983. An overview of the PENMAN text generation system. In *Proceedings of the National Conference on Artificial Intelligence*, 261–265. AAAI. Also appears as USC/Information Sciences Institute, RR-83-114.

Mann, William C. & Christian M. I. M. Matthiessen. 1985. Demonstration of the Nigel Text Generation Computer Program. In James D. Benson & William S. Greaves (eds.), *Systemic Perspectives on Discourse, Volume 1*, 50–83. Norwood, New Jersey: Ablex.

Manning, Christopher D, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J Bethard & David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. In *Acl 2014*, 55–60. <http://www.aclweb.org/anthology/P/P14/P14-5010>.

Marneffe, Marie-Catherine, Timothy Dozat, Natalia Silveira, Katri Haverinen, Filip Ginter, Joakim Nivre & Christopher D. Manning. 2014. Universal Stanford dependencies: A cross-linguistic typology. In *Proceedings of the ninth international conference on language resources and evaluation (lrec-2014)(vol. 14)*, Reykjavik, Iceland: European Language Resources Association (ELRA). <http://www.lrec-conf.org/proceedings/lrec2014/pdf/1062_Paper.pdf>.

Marneffe, Marie-Catherine, Bill MacCartney & Christopher D Manning. 2006. Generating Typed Dependency Parses from Phrase Structure Parses. In *Lrec 2006*, vol. 6 3, 449–454. Stanford University. <http://nlp.stanford.edu/manning/papers/LREC_2.pdf>.

Matthiessen, Christian M. I. M. 2015a. Halliday's Conception of Language as a Probabilistic System. In Jonathan J Webster (ed.), *The Bloomsbury Companion to M.A.K. Halliday*, 203–241. London and New York: Bloomsbury.

Matthiessen, Christian M.I.M. 2015b. Register in the round: registerial cartography. *Functional linguistics* 2(9). 1–48. doi:10.1186/s40554-015-0015-.

Matthiessen, M.I.M., Christian. 1985. The systemic framework in text generation: Nigel. In James Benson & Willian Greaves (eds.), *Systemic perspective on Discourse, Vol I*, 96–118. Norwood, New Jersey: Ablex.

McDonald, Daniel & Robyn Woodward-Kron. 2016. Member roles and identities in online support groups: Perspectives from corpus and systemic functional linguistics. *Discourse & Communication* 10(2). 157–175.

McDonald, David D. 1980. *Natural Language Production as a Process of Decision Making under Constraint*: MIT, Cambridge, Mass dissertation.

Neale, Amy C. 2002. More Delicate TRANSITIVITY: Extending the PROCESS TYPE for English to include full semantic classifications. Tech. rep. Cardiff University Cardiff.

Nivre, Joakim. 2015. Towards a universal grammar for natural language processing. In *Computational Linguistics and Intelligent Text Processing*, 3–16. Berlin: Springer.

O'Donnell, Michael. 1990. A dynamic model of exchange. *Word* 41(3). 293–328.

O'Donnell, Michael. 1993. Reducing Complexity in Systemic Parser. In *Proceeedings of the third international workshop on parsing technologies*, Tilburg.

O'Donnell, Michael. 1994. *Sentence analysis and generation: a systemic perspective*. Sydney, Australia: University of Sydney, Department of Linguistics dissertation. < http://www.wagsoft.com/Papers/Thesis/index.html>.

O'Donnell, Michael. 2005. The UAM Systemic Parser. In *Proceedings of the 1st Computational Systemic Functional Grammar Conference*, University of Sydney, Sydney, Australia. < http://www.wagsoft.com/Papers/ODonnellUamParser.pdf>.

O'Donnell, Michael. 2008. Demonstration of the UAM corpus tool for text and image annotation. In *Proceedings of the ACL '08: HLT Demo Session*, 13–16. Columbus, Ohio: Association for Computational Linguistics. < http://www.aclweb.org/anthology-new/P/P08/P08-4004.pdf>.

O'Donnell, Michael J. & John A. Bateman. 2005. SFL in computational contexts: a contemporary history. In Ruqaiya Hasan, Christian M.I.M. Matthiessen & Jonathan J. Webster (eds.), *Continuing Discourse on Language: A functional perspective*, vol. 1, 343–382. London and New York: Equinox.

Oepen, Stephan, Dan Flickinger, Hans Uszkoreit & Jun ichi Tsujii. 2000. Introduction to the special issue on efficient processing with HPSG. *Natural Language Engineering* 6(1). 1–14.

O'Halloran, Kay L. 2008. Systemic Functional-Multimodal Discourse Analysis (SF-MDA): Constructing Ideational Meaning using Language and Visual Imagery. *Visual Communication* 7(4). 443–475.

O'Halloran, Kay L. 2014. Multimodal Discourse Analysis. In Ken Hyland & Brian Paltridge (eds.), *The Bloomsbury Companion to Discourse Analayis*, 120–137. London and New York: Bloomsbury.

O'Halloran, Kay L. 2015. Multimodal Digital Humanities. In Peter P. Trifonas (ed.), *International handbook of semiotics*, 389–416. Dordrecht: Springer.

O'Halloran, Kay L., Alvin Chua & Alexey Podlasov. 2014a. The role of images in social media analytics: A multimodal digital humanities approach. In David Machin (ed.), *Visual communication*, 565–588. Berlin: Mouton de Gruyter.

O'Halloran, Kay L., Marissa K.L.E & Sabine Tan. 2014b. Multimodal analytics: software and visualization techniques for analyzing and interpreting multimodal data. In Carey Jewitt (ed.), *The Routledge Handbook of multimodal analysis*, 386–396. London: Routledge 2nd edn.

O'Halloran, Kay L., Sabine Tan, Duc-Son Pham, John A. Bateman & Andrew Vande Moere. 2016a. A Digital Mixed Methods Research Design: Integrating Multimodal Analysis with Data Mining and Information Visualization for Big Data Analytics. *Journal of Mixed Methods Research* doi:10.1177/1558689816651015.

O'Halloran, Kay L., Sabine Tan, Peter Wignell, John A. Bateman, Duc-Son Pham, Michele Grossman & Andrew Vande Moere. 2016b. Interpreting text and image relations in violent extremist discourse: A mixed methods approach for big data analytics. *Terrorism and Political Violence* doi:10.1080/09546553.2016.1233871.

O'Toole, Michael. 2011. *The language of displayed art*. Abingdon, Oxon: Routledge 2nd edn.

Peters, Stanley P. & Richard W. Ritchie. 1973. On the generative power of transformational grammars. *Information Sciences* 6. 49–83.

Rautaray, Siddharth S. & Anupam Agrawal. 2015. Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review* 43(1). 1–54. doi:10.1007/s10462-012-9356-9.

Richardson, Matthew & P. Domingos. 2006. Markov logic networks. *Machine learning* 62(1-2). 107–136. doi:10.1007/s10994-006-5833-1. < http://link.springer.com/article/10.1007/s10994-006-5833-1>.

Schreibman, Susan, Ray Siemens & John Unsworth (eds.). 2016. *A new companion to digital humanities*. Chichester: Wiley-Blackwell 2nd edn.

Shi, Hui, Cui Jian & Carsten Rachuy. 2011. Evaluation of a Unified Dialogue Model for Human-Computer Interaction. *International Journal of Computational Linguistics and Applications*

2(1). 155–173.

Sleator, Daniel D. K. & Davy Temperley. 1995. Parsing English with a link grammar. *arXiv preprint cmp-lg/9508004* .

Socher, Richard, John Bauer, Christopher D Manning & Andrew Y Ng. 2013. Parsing with Compositional Vector Grammars. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, 455–465.

Steedman, Mark J. 1993. Categorial Grammar. *Lingua* 90. 221–258.

Steels, Luc. 2005. The emergence and evolution of linguistic structure: from lexical to grammatical communication systems. *Connection Science* 17(3–4). 213–230.

Svensson, Patrick. 2010. The landscape of digital humanities. *Digital Humanities Quarterly* 4(1). < http://www.digitalhumanities.org/dhq/vol/4/1/000080/000080.html>.

Teich, Elke, Stefania Degaetano-Ortlieb, Peter Fankhauser, Hannah Kermes & Ekaterina Lapshinova-Koltunski. 2016. The Linguistic Construal of Disciplinarity: A Data Mining Approach Using Register Features. *Journal of the Association for Information Science and Technology (JASIST)* 67(7). 1668–1678.

Teich, Elke, Eli Hagen, Brigitte Grote & John A. Bateman. 1997. From communicative context to speech: integrating dialogue processing, speech production and natural language generation. *Speech Communication* 21(1–2). 73–99.

Traum, David & Staffan Larsson. 2003. The Information State Approach to Dialogue Management. In Ronnie Smith & Jan van Kuppevelt (eds.), *Current and New Directions in Discourse and Dialogue*, 325–353. Dordrecht: Kluwer Academic Publishers.

Tseng, Chiao-I. 2013. *Cohesion in Film: Tracking Film Elements*. Basingstoke: Palgrave Macmillan.

Weerasinghe, Ruvan. 1994. *Probabilistic Parsing in Systemic Functional Grammar*: University of Wales College of Cardiff dissertation.

Winograd, Terry. 1972. *Understanding Natural Language*. New York: Academic Press.

XTAG Research Group. 2001. A Lexicalized Tree Adjoining Grammar for English. Tech. Rep. IRCS-01-03 IRCS, University of Pennsylvania. < http://www.cis.upenn.edu/x̃tag/tech-report/tech-report.html>.

Zinn, Jens O. & Daniel McDonald. 2015. Changing Discourses of Risk and Health Risk. In John M. Chamberlain (ed.), *Medicine, risk, discourse and power*, 207–240. Routledge.