ΣynΣem

From Form to Meaning: Integrating Linguistics and Computing

— Proposal to the Oslo Center for Advanced Studies —
(Version of March 21, 2015; SVN Revision 12713)

Dag Trygve Truslew Haug
Department of Philosophy, Classics, History of Arts and Ideas
University of Oslo
d.t.t.haug@ifikk.uio.no

Stephan Oepen
Department of Informatics
University of Oslo
oe@ifi.uio.no

Abstract

Computational linguistics investigates human language through computational techniques. Our language processing capacity is at the core of human intelligence; language provides the predominant channel of inter-human communication; and digitally encoded language is universally recognized as the 'fabric' of the World-Wide Web. Thus, computational linguistics is of increasing societal relevance, culturally as well as economically. The computational investigation of human language is an inherently inter-disciplinary field, with a direct bearing on both the humanities (notably linguistics and philosophy) and the sciences (mathematics and computing).

However, a growing number of successful applications of computational linguistics and a related increased interest in practical, engineering approaches have led to a partitioning of research in computational linguistics—into either predominantly theoretical or primarily applied perspectives. Adverse effects of this dichotomy are clearly visible today: methodological fragmentation, on the one hand, and plateau effects in engineering progress, on the other hand. Our proposal for an intensive, one-year, multi-disciplinary research group will help reduce fragmentation, re-unite and henceforward advance in tandem historically related sub-disciplines, as well as prepare the field (and participating researchers) for emerging and future challenges in the formal, computational, and applied analysis of human language.

As a defining starting assumption, shared among the proposers, our initiative pursues a holistic, long-term perspective, seeking to bring into equilibrium theoretical as well as practical desiderata: Such integration, we believe, will be a prerequisite to both advancing our understanding of the human language capacity and making core insights from linguistics applicable (and accessible) to the design of next-generation, language-enabled computational systems, affecting inter-human communication and human–computer interaction alike.

NORSK SAMMENDRAG

Datalingvistikken utforsker menneskets språk med hjelp av datamaskiner. Språkevnen er i kjernen av våre intellektuelle ferdigheter og språket er vårt dominerende kommunikasjonsmiddel. Stadig mer informasjon er tilgjengelig som digital tekst på internett. Datalingvistikken er derfor av stadig større betydning for samfunnet, både kulturelt og økonomisk.

Å forske på språk med komputasjonelle metoder er et tverrfaglig felt som spenner over både humaniora (særlig lingvistikk og filosofi) og naturfagene (matematikk og informatikk). Slik forskning kan ha et rent teoretisk perspektiv, hvor man f.eks. forsøker å formulere generaliseringer om grammatikken og fange språket som et formelt system, gjerne på en måte som i prinsippet lar seg implementere på en datamaskin, men uten direkte tanke for slike anvendelser. Eller man kan ha et rent anvendt perspektiv, hvor man forsøker å prosessere tekster med en datamaskin uten å ta sikte på en dypere forståelse av hvordan språk fungerer.

I en ideell verden ville teoretiske framskritt føre til bedre anvendelser, men den suksessen applikasjonsorienterte metoder har hatt i det siste har faktisk ført til en kløft mellom anvendt og teoretisk forskning. Anvendelsene bruker gjerne 'grunne' statistiske metoder som 'lærer' kategoriseringer direkte fra data. Dette fungerer godt i mange applikasjoner (tenk f.eks. på Google Translate), men det er tegn som tyder på at man nå støter mot grenser for hvor langt man kan nå med metoder som ikke bygger på teoretisk forståelse av språk.

Målet for vårt CAS-prosjekt er å minske fragmenteringen i feltet og bygge bro mellom de forskjellige tilnærmingene. Vi har et holistisk perspektiv og mener at en integrasjon av teoretiske og anvendte perspektiver vil være til gjensidig nytte, og nødvendig både for å oppnå en bedre forståelse av den menneskelige språkevnen og for å utvikle applikasjoner som vil sette den neste generasjonen datamaskiner bedre i stand til å forstå språk. Et grunnleggende premiss for vårt prosjekt er at de ulike metodene, teoriene og teknikkene som er i bruk nå er modne nok til at det er mulig å kombinere og raffinere dem i et helhetlig perspektiv.

Automatisert språkforståelse forutsetter en algoritme som lar oss gå fra rå tekst via en strukturanalyse av setningene til en representasjon av språklig betydning. Prosjektet fokuserer på tre problemområder som står sentralt i en slik prosess: Hvordan kan vi best representere setningers grammatiske form? Hvordan kan vi kople ordenes grammatiske funksjon til deres semantiske rolle i setningen? Og hvordan kan vi bygge opp representasjoner av en setnings betydning fra betydningen av ordene i den?

Studiet av setningers grammatiske form dekkes av *syntaks*, som bygger på metoder fra det som i informatikken heter formell språkteori. Det finnes flere skoler innen syntaktisk formalisering og vårt prosjekt har deltakere fra to av de teoriene som tradisjonelt har fokusert på å bygge bro mellom lingvistisk teori og komputasjonell anvendelse, nemlig *Head-Driven Phrase Structure Grammar* (HPSG) og *Lexical-Functional Grammar* (LFG). Skjønt det er mange likheter i mål og oppbygning mellom disse teoriene, så er det også mange forskjeller. Et hovedmål for prosjektet er å samle forskere fra disse ulike tradisjonene under ett tak, for å utveksle erfaringer og komme videre både innen det teoretiske studiet av syntaks og i den praktiske anvendelsen av syntaktisk teori.

Et viktig mål både i lingvistisk teori og i anvendte systemer som tar sikte på å få datamaskiner til å 'forstå' språk, er å bygge semantiske representasjoner fra de syntaktiske analysene. Her har teoretisk lingvistikk tradisjonelt fokusert på *komposisjonalitet*, hvordan setningsbetydningen bygges opp, men har brukt forenklede representasjoner av 'byggesteinene', enkeltordenes betydning. Teknologiske anvendelser har på sin side fokusert på datadrevne representasjoner av ordbetydning, uten særlig tanke på komposisjonalitet. Det er mange utfordringer ved å bruke komposisjonelle teknikker fra teoretisk lingvistikk på slike betydningsrepresentasjoner, og dette er for tiden et aktivt forskningsfelt med viktige bidrag fra deltakere i vår kjernegruppe.

1 Introduction

In the early 1950s, COMPUTATIONAL LINGUISTICS emerged as a new discipline, growing out of both a *cognitive-science vision* and a *practical demand*. The bold vision, on the one hand, was to apply emerging computational techniques to the understanding of human language. The engineering need, on the other hand, with roots in war-time code cracking, responded to a strongly felt demand for automated translation. Scholars in linguistics, philosophy, mathematics, computer science, and psychology jointly set out to investigate the system of language, to formally analyze the *structure of linguistic utterances*, and to model their *patterns of use*. New computational tools and a gradual development towards the 'digitization' of language made it possible to study ever bigger language samples, using both *qualitative* and *quantitative* methods. Mathematically precise modeling made it possible to test theories formally, thus putting linguistics onto a solid formal footing.

Some six decades later, formal theories of grammar have been developed and are being applied to a broad range of languages. There are large, digital collections of data from diverse languages, some annotated with linguistic analyses. The engineering branch of computational linguistics has matured into a sub-discipline of computer science—known as (Human) LANGUAGE TECHNOLOGIES—that helps develop services like the Microsoft Grammar Checker, Google Translate, or Apple's Siri. However—maybe following early disillusionment about short-term results, maybe out of scientific necessity—the many sub-disciplines of computational linguistics have evolved largely independently and 'immunized' themselves to some degree. The models developed in formal syntax and semantics are rarely applied directly in the creation of linguistically interpreted corpora, i.e. annotations of structure and meaning; likewise, many mainstream techniques in applied language technologies are uninformed by theoretical results in formal linguistics, and vice versa.

Fernando Pereira, one of the pioneers in computational linguistics, serves as Research Director at Google today; in the abstract to a 2013 keynote, he noted:

> Advances in statistical and machine learning approaches to natural-language analysis have yielded a wealth of methods and applications in information retrieval, speech recognition, machine translation, and information extraction. Yet, even as we enjoy these advances, we recognize that our successes are to a large extent the result of clever exploitation of redundancy in language structure and use, allowing our algorithms to eke out a few useful bits that we can put to work in applications. By focusing on applications that extract a limited amount of information from the text, finer structures such as word order or syntactic structure could be largely ignored [...]. However, [...] our language-processing systems have been stuck in an 'idiot savant' stage where they can find everything but cannot understand anything. [...] I will argue with examples from our recent research that we need deeper levels of linguistic analysis to do this. (`http://www.clsp.jhu.edu/seminars/1422/`)

2 Vision and Hypotheses

Our vision for the stay at the Center for Advanced Studies (CAS) is to reduce fragmentation in computational linguistics—by taking advantage of the intense and sheltered cooperation made possible only in a setting like this. Thus, our ambition is *integrational*, to build bridges across entrenched lines of division, to take stock in the state of the art in related sub-disciplines, to establish and strengthen points of contact and overlapping interests, and to prepare a push towards *unifying models and techniques*. As argued below, this is a timely vision for the field at large, as it emboldens a recent undercurrent of disillusionment and desire for re-balancing, particularly as a whole generation of leading figures is approaching retirement. It is also an appropriate ambition for the proposers, who are well equipped to advance these developments—by virtue of their complementary specializations, career stages, participation in relevant international networks, and inter-faculty institutional support at Oslo University.

Our proposal is rooted in two key hypotheses, viz.

(a) that theories, methods, and techniques in the various disciplines contributing to computational linguistics have *matured* to a point that makes it possible to 'shake and bake', i.e. combine and refine in a strongly integrational, holistic perspective; and

(b) that such a *holistic approach*—drawing on formal, empirical, and engineering expertise—is a prerequisite to overcoming a trend of diminishing returns in the recent past, i.e. what practitioners in language technologies perceive as plateau or ceiling effects on engineering progress.

The past decade has seen great advances in computational approaches to natural languages. In linguistics, lexicography, the philologies, and other areas of the humanities, much research today builds on digitized language corpora and techniques for search, aggregation, or automated hypothesis testing. These developments have enabled a stronger empirical grounding of language studies. At the same time, high-visibility applications of computational linguistics have entered the consumer market, for example in speech recognition, automated translation, and human–computer dialogue; and there are numerous less visible such applications at play in monetizing on-line content. Despite undeniable commercial success, current technologies fall far short of human linguistic capacities, however: Google Translate is not reliable enough to allow one to enter a legal contract in a foreign language, say; and any non-trivial dialogue with Apple's Siri is more likely to be entertaining than helpful. These applications do not directly advance our understanding of human language either, as long as they predominantly build on techniques that are theory- and knowledge-poor, i.e. remain disconnected from contemporary linguistic research.

The proposed research group unites long-term, informally organized networks in computational linguistics with several more recent initiatives in the realm of linguistically informed practical language technologies. These are, on the one hand, the Parallel Grammars Project (ParGram; `http://pargram.b.uib.no`), based on the linguistic theory of Lexical Functional Grammar (LFG; Kaplan & Bresnan, 1982; Bresnan, 2001) and the Deep Linguistic Processing with HPSG Initiative (DELPH-IN; `http://www.delph-in.net`), based on Head-Driven Phrase Structure Grammar (HPSG; Pollard & Sag, 1987, 1994). The proposers have long been active members of these networks: Haug has participated in the LFG community and been associated with ParGram since 2008, and Oepen was one of the DELPH-IN co-founders in the late 1990s. Other participants in the proposed group represent the Universal Dependencies (UD) initiative, a new multi-national effort towards cross-linguistically consistent syntactic annotation (de Marneffe et al., 2014; Nivre, 2015), as well as ongoing work on large-scale annotation and processing of dependency-based semantic representations, including in the frameworks of the Prague Functional Generative Description (FGD; Sgall et al., 1986; Hajič et al., 2012) and Abstract Meaning Representation (AMR; Banarescu et al., 2013). The group will bring together, for the first time, key international players representing different perspectives from within these paradigms, which historically have developed independently or even competitively. Participants jointly emphasize common goals and share a vision of re-energizing cross-fertilization between these communities.

3 Background: Dimensions of Variation

We envision a cross-disciplinary research group that integrates heterogeneous but kindred specializations in theoretical and computational linguistics. In the following, we identify three foundational axes of variation that jointly set up a multi-dimensional architecture—and points of contact—for our research programme. For each dimension, we provide a high-level assessment of the state of the art and current and anticipated obstacles to scientific progress. In Section 5 below, we then define two in-depth research challenges for our group that each span across several dimensions of variation.

3.1 The Relationship Between Form and Meaning

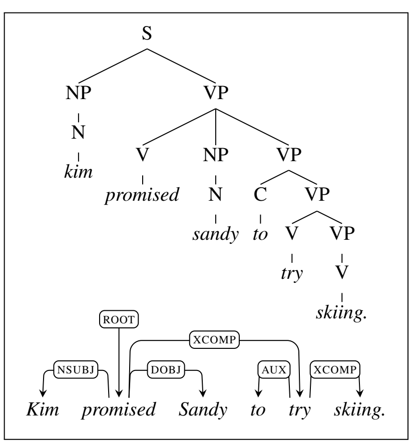

Figure 1: Candidate analyses for (1).

Mapping between SENTENCE FORM (syntax) and SENTENCE MEANING (semantics) takes a central role in theoretical linguistics, with the traditional assumption that semantic composition is guided by syntactic structure. Syntactic representations can be coarsely sub-divided into constituency- vs. dependency-based ones, as depicted in Figure 1 for the running example:

(1) *Kim promised Sandy to try skiing.*

Both *promise* and *try* are lexically classified as subject-equi verbs, such that the unexpressed subjects of their non-finite verbal complements are controlled by the matrix subject (Sag & Pollard, 1991). Whatever formalism one adopts, the ultimate goal for grammar theory (as well as for applications in the pursuit of natural language 'understanding') is to map syntactic structures to a representation of *meaning*, typically in a suitably abstract and mathematically interpretable form. In theoretical linguistics and philosophy, this mapping is assumed to instantiate the PRINCIPLE OF COMPOSITIONALITY commonly attributed to logician Gottlob Frege. In combination with the lexical categorization of the two subject-equi verbs in (1), both the constituent tree (top in Figure 1) and the bi-lexical dependencies (bottom) encode the necessary syntagmatic relations to let (the semantics associated with) *Kim* bind the semantic argument roles of the two controlled verbs. How exactly to establish this binding, however, is a foundational question at the syntax–semantics interface, and choices made at this level have important formal and processing ramifications.

Regarding design choices for representations of meaning, generally speaking, much theoretical formal semantics has focused on the *combination* of meanings, but paid little attention to the basic building blocks. Lexical semantics, on the other hand, has investigated *word meanings*, which one can either seek to decompose in terms of semantic primitives (see Pustejovsky, 1991, for classic references) or—in much current work—approximate by the study of word *distributions* over contexts of use in large corpora. Obviously, the two perspectives must work hand in hand in a comprehensive account of natural language meaning, and there is growing interest in reconciling compositional and distributional approaches to semantics (Lewis & Steedman, 2013; Baroni et al., 2014; inter alios; see Section 5.2 below).

LFG and HPSG—each with active researcher communities world-wide and a strong representation among core members of our research group—are two theories that since their inception in the early 1980s have paid equal attention to formal and computational considerations, and that by design have taken a holistic perspective—encompassing morpho-syntax as well as semantics. Hence, they have enabled mature computational implementations for a broad variety of languages. Besides much abstract similarity in foundations and goals, there are several TECHNICAL DIVERGENCES between these frameworks, and these hinder cross-fertilization and impact. Bringing together researchers from the multi-national ParGram and DELPH-IN networks described above for a full year at CAS will create a unique opportunity to (a) exploit areas of commonality and overlap between these schools and to (b) relate key insights and mechanisms at their syntax–semantics interface to the more recent UD and AMR initiatives, which currently are focused near-exclusively on either syntactic analysis or meaning representation, respectively.

3.2 Categorical vs. Statistical Regularities

Non-categorical, GRADED DISTRIBUTIONS in language (or 'soft' constraints or preferences)—for example in judging grammaticality—challenge the 'classic' formal models of grammar sketched above. At the same time, syntactico-semantic analysis by itself is insufficient for most applied contexts: Computational language 'understanding' also requires disambiguation, i.e. ranking candidate analyses in terms of plausibility. These observations (and others, for example in the study of language acquisition) call for the integration of FREQUENCIES (and probabilities) as first-class elements in theoretical as well as applied computational linguistics.

Practical language technologies have embraced statistical regularities (and, correspondingly, machine learning theory), but as suggested in the reflections by Pereira above this development has arguably weakened their grounding in formal linguistics. With the availability of very large digital language samples, advanced mathematical models of probability distributions, and greatly increased computing resources have come some 'break-through' results that for a while at least appeared to prepare the grounds for a 'statistical sea change' in computational linguistics. However, in recent history it is (again) acknowledged that (a) a non-trivial part of natural language structures, to say the least, cannot be learned exclusively as emergent properties from 'raw' language data; and that (b) the vast complexity, diversity, and adaptability of language together can quickly give rise to an intractable sample space on current or mid-term computing resources. Therefore, also predominantly statistical approaches to language analysis are increasingly looking to use linguistic generalizations that help structure that space better.

The division of labor between categorical linguistic rules and statistical regularities related to frequencies of use remains an acute research question. Through integration in our group of scholars who specialize in statistical, data-driven syntactic and semantic analysis, we seek to probe this question through in-depth contrastive studies and, thus, to help reconcile 'analytical' and 'empirical' modeling.

3.3 From Theory via Annotation to Processing

Finally, our third dimension of variation and source of tension between work in theoretical and (applied) computational linguistics ranges from formal syntax and semantics, on the one hand, over empirical, descriptive linguistics, on the other hand, and all the way to applied language technologies. Along this axis, linguists focus on different short-term goals and apply only partly overlapping methodologies. In the abstract, linguistic theory should ideally translate directly into annotation best practices, i.e. pairing natural language samples with gold standard syntactico-semantic analyses. Likewise, linguistically principled, machine-readable annotation would feed directly into computational processing, be it as the empirical basis for analytical grammar engineering or as training data for statistical learning.

In practice, however, the relationships between THEORY, ANNOTATION, and PROCESSING are not at all straightforward, given the theoretical fragmentation observed earlier, the cost of manual annotation, the need to apply annotation to diverse languages and graded linguistic phenomena, and technical requirements imposed by various computational processing tasks. In the context of our research group, our processing focus will be on automated parsing, i.e. mapping from linguistic signals to abstract representations of structure and meaning. But research on natural language parsing in a computer science perspective, today, oftentimes lacks access to core linguistic traditions.

In the ParGram and DELPH-IN networks, linguistic theory and automated analysis are synchronized through computational grammars, i.e. declarative knowledge about language regularities, that are combined with probabilistic models for disambiguation. Our research group complements these competencies with strong expertise in (a) large-scale syntactico-semantic annotation (of naturally occuring text in multiple languages) and in (b) data-driven parsing, i.e. the design of statistical algorithms that acquire linguistic knowledge and preferences of use from annotated data. In this regard, much work in data-driven dependency parsing can be characterized as 'grammar-less', i.e. in principle recognizing arbitrary strings and shifting the complete analysis task into disambiguation. Nevertheless, the UD initiative—combining annotation creators and parser developers—historically derives some of its core assumptions from LFG, and extension of UD to more languages calls for cross-linguistic grounding in current syntactic research. Also, it is plausible to expect that going from syntactic dependencies to representations of meaning (like those of AMR, say) will benefit from explicit and formal models of semantic composition.

4 The Group

This proposal is rooted in a new inter-faculty initiative at the University of Oslo, where researchers in the Faculty of Humanities and in the Department of Informatics have joined forces to collaboratively investigate the theoretical and applied relations between language form and meaning. This initiative, dubbed SYNSEM, is headed by the proposers together with professor Atle Grønn (`http://folk.uio.no/atleg/`) and associate professor Lilja Øvrelid (`http://folk.uio.no/liljao/`). SYNSEM has hired two post-doctoral and three doctoral fellows with internal seed funding from the two faculties. These early-stage researchers are taking up their positions in early 2015 and will remain associated to SYNSEM during our proposed CAS stay. We will reserve one of the offices at CAS for SYNSEM early-stage researchers.

Against this emerging initiative, the opportunity to collaborate intensely with leading international experts for a full year at CAS will have a strong and immediate positive effect on the academic and organizational cohesion of the SYNSEM group, promising to forge a cross-disciplinary and multi-generational community of high international standing and visibility. Besides the proposers and their SYNSEM collaborators, four leading scholars, introduced briefly in the Appendix, have agreed to participating in the group for periods of between six and twelve months. Another eight top-tier international scholars, also identified in the Appendix, plan on participating in the group through extended visits on the order of several weeks to several months.

5 Research Challenges

Within the three-dimensional grid established by the above linguistic and methodological axes, we aim to establish collaboration across different sub-disciplines and, ideally, mutual cross-fertilization by organizing the research group internally according to strands of investigation that cut across multiple dimensions. The challenges concern both syntax, semantics and the interface between the two. As outlined in more detail in Section 6 below, we take an empirical, phenomenon-driven approach. In the following sections we discuss how the phenomenon of grammatical control exemplifies these challenges, thus providing a running 'case study' for the proposed mode of operation of our CAS group. But we conclude each section with a set of 'big-picture' research questions that cut across invididual phenomena, as the final inventory of linguistic phenomena shall be determined collectively at the start of our period at CAS.

A source of variation inherent in any study of human language is LINGUISTIC DIVERSITY, i.e. substantial differences in the grammatical systems used to relate form to meaning. Even more so than in theoretical linguistics, much work in computational linguistics has focused on English (Bender, 2011).

And although standard models often are proposed as general, language-independent frameworks, their application to different languages typically calls for extension or refinement. Besides English, Scandinavian, and other Germanic languages, our research group encompasses strong competencies on a variety of language families, including Australian, Austronesian, Greek, Japonic, Latin and Slavic.

5.1 Organization of Grammatical Information

The distinction between emphasizing constituent vs. dependency structure—as evidenced in the two candidate syntactic analyses of Figure 1 above—is at the core of much disagreement in modern syntactic theory. Traditionally, following Chomsky (1957), 'mainstream' Western formal grammar assumed a primacy of constituency. This tradition is largely continued in HPSG, while in LFG dependencies (known as functional structures) play a fundamental role in grammatical theory. Dependencies can be argued to be 'closer' to underlying semantic relations, i.e. what is at times called *Who did What to Whom?* (Nivre, 2006). On the other hand, the notion of constituency makes available additional layers of hierarchical structure—for example a distinction between verbs, verb phrases, and complete sentences—which may have utility at the syntax–semantic interface.

$$\left\{\begin{array}{l}\text{promise\_1}(\text{ARG0 } e_1, \text{ ARG1 } x_1, \text{ ARG2 } x_2, \text{ ARG3 } h_1),\\ \text{named}(\text{ARG0 } x_1, \text{ CARG } \textit{Kim}),\\ \text{named}(\text{ARG0 } x_2, \text{ CARG } \textit{Sandy}),\\ h_1\text{:try\_1}(\text{ARG0 } e_2, \text{ ARG1 } x_1, \text{ ARG2 } h_2),\\ h_2\text{:ski\_1}(\text{ARG0 } e_3, \text{ ARG1 } x_1)\end{array}\right\}$$

Figure 2: Predicate–argument relations in (1).

Going back to our running example (1), it is uncontroversial to expect the syntax–semantics interface to equate the 'agent'(ive) roles of all predications introduced by verbs. Figure 2 sketches a candidate interface representation for grammatical analysis, using the notation of 'flat' Minimal Recursion Semantics (MRS; Copestake et al., 2005; though formal details are immaterial here). In this analysis, there are two entities x_1 and x_2 (named *Kim* and *Sandy*, respectively) that participate in Davidsonian eventualities $e_1 \ldots e_3$; the labeling of individual predications (e.g. h_1 for the semantic contribution of *try*) and use of labels in argument positions effectively sub-ordinate the fragment rooted in h_1 as the third argument of *promise*. The reflection of obligatory control in this example, as would be expected, is the co-indexation of entity variable x_1 as the ARG1 in all three eventualities.

Both the HPSG and LFG theories provide integrated, declarative mechanisms for semantic composition guided by syntactic structure: HPSG uniformly combines morpho-syntactic and semantic information in its mono-stratal *sign* architecture (including a semantic representation like the one in Figure 2); conversely, the LFG projection architecture pairs the derivation of a constituent tree like that of Figure 1 (top) with the evaluation of functional equations (yielding structures as in Figure 3 below). In both cases, syntactico-semantic information is formally encoded in feature structures and manipulated via graph unification (Shieber, 1986; Carpenter, 1992; inter alios), which has given rise to the genus term *unification-based grammars*. The fine-grained recursive structure of linguistic information in these frameworks, combined with the power of graph unification, makes it comparatively straightforward to give a concise and declarative account of common sub-problems in semantic composition, including for example linking grammatical functions to thematic argument positions, scopal embedding, and intersective modification. At the same time, the HPSG and LFG formalisms exceed the class of mildly context-sensitive languages, i.e. they align poorly the generative capacity of the description language with the actual structural complexity of natural languages.

Dependency grammar constitutes a more heterogeneous family of theories that share a common set of assumptions. The fundamental hypothesis in all versions of dependency grammar is that the core of syntactic structure can be captured by binary asymmetric relations between lexical units, as illustrated above in Figure 1 (bottom). Some frameworks, like Meaning–Text Theory (Mel'čuk, 1988), extend the dependency-based view also to morphology and semantics, and it is common in the dependency grammar tradition to posit two levels of dependency structure, a surface syntactic structure that provides the interface to morphology, and a deep syntactic structure that provides the interface to semantics (Sgall et al., 1986; Mel'čuk, 1988; Hajič et al., 2012). These syntactic representations are normally assumed to be trees, but an alternative to using multi-stratal representations is to use a richer dependency structure that is no longer a tree but a general graph. The latter approach can be found (informally) in the Structural Syntax of Tesnière (1959) and in the framework of Word Grammar by Hudson (1984).[1]

[1] These frameworks also differ from the mainstream tradition by not trying to reduce all grammatical constructions to dependency relations. Notably, coordination is treated as a fundamentally different syntactic construction.

Despite the rich variety of representations found in dependency grammar, most computational implementations—especially in statistical parsing—currently limit themselves to a single *tree*, which can represent either surface or deep syntax, or a mixture of both. Applied to (1) above, for example, parsing into fully connected trees (where each node is dependent on exactly one head) can be linguistically problematic: Under this formal constraint, a dependency analysis as in Figure 1 cannot capture the control of unexpressed subjects, nor could the infinitival particle *to* be analyzed as vacuous in a perspective shifting emphasis towards 'deep' syntax or predicate–argument structure. The recent Universal Dependencies initiative creates representations that go beyond a basic dependency tree in an attempt to incorporate all relations that are required by the syntax–semantic interface, while still supporting efficient processing by statistical dependency parsers (Nivre, 2015).

In a parallel development, there is growing interest in the use of dependency-based general *graphs* for meaning representation, removing the constraints on connectedness and single-headedness. Moving from tree- to graph-structured target representations in statistical parsing presents no small challenge, as models and algorithms until recently were near-exclusively developed for the formally much simpler tree-centered view. At the same time, there are no worked out mechanisms for algebraic meaning composition guided by the simple syntactic dependency trees of Figure 1, and current best practice is to let a statistical semantic parser be 'weakly' informed by syntax through dependency indicator features. Two shared tasks at recent Semantic Evaluation Exercises (SemEval), co-organized by Stephan Oepen, Dan Flickinger, and Jan Hajič, have generated initial momentum in the dependency parsing community by reducing native representations of sentence meaning from the HPSG and the Prague FGD frameworks to bi-lexical semantic dependency graphs (Oepen et al., 2014, 2015).[2] In this context, there is an acute need for research that relates linguistic and formal properties of the target representations, on the one hand, with computational and mathematical aspects of the emerging processing models, on the other hand.

Graph-structured representations of (contextually interpreted) speaker meaning, as recently pioneered by the AMR initiative (Banarescu et al., 2013; Xue et al., 2014), are increasingly considered as candidate target representations for data-driven parsing (Flanigan et al., 2014). AMR integrates layers of interpretation that are not traditionally encompassed by grammatical analysis (e.g. some word sense disambiguation and resolution of coreference) and, thus, its design raises foundational questions on the role of grammar, notions of compositionality, and the interplay between formal and computational semantics and more general language 'understanding' (Bender et al., 2015).

Established linguistic tradition seeks to delineate sub-components and define interfaces (e.g. between morphology, syntax, and semantics), which in HPSG and LFG are established within the mono-stratal or the projection architectures, respectively. Similarly, much current work in computational language analysis pursues so-called joint (tightly integrated) modeling—in part in order to reduce system-internal complexity, in part seeking to eliminate the propagation of errors through separate components in a pipeline archicture. However, most statistical parsing work in this direction is uninformed (likely often unaware, in fact) of the fully worked out mechanisms for (a) integrating syntax and semantics and (b) graph-structured target representations in unification-based theories. Therefore, it is most timely for the proposed research group to pursue the in-depth contrastive comparison of approaches—in terms of linguistic as well as computational properties.

The design and choice of syntactico-semantic representations is not just of theoretical and computational significance; it is also central in the annotation of language samples with gold-standard analyses (i.e. 'treebanks'). Here there is a practical trade-off between the desire to capture as much as possible of the grammatical structure of a sentence, which may be theoretically disputed, and the need for efficient and consistent annotation at scale. Existing treebanks, of constituent or dependency structure alike, reflect many simplifications that are pragmatically rather than theoretically motivated. Accordingly, current data-driven syntactic parsers remain limited by the style and granularity of available annotations and, thus, deliver analyses that at times have little in common with the types of structures a theoretical linguist would posit. More recently, some treebanks are being created providing comparatively 'deeper' analyses within the theoretical frameworks (and building on the computational grammars) represented in our group (King et al., 2003; Oepen et al., 2004; Flickinger, Zhang, & Kordoni, 2012; Basile et al., 2012; inter alios). A first series of contrastive parser evaluations (by 'reduction' of HPSG analyses to syntactic dependencies) suggests that richer syntactic representations can lead to improved parser accu-

[2]The shared task web sites remain available at `http://alt.qcri.org/semeval2014/task8/` and `http://alt.qcri.org/semeval2015/task18/`.

racy and greater resiliance to variation in text types (Plank & van Noord, 2010; Ivanova et al., 2012, 2013). Furthermore, 'deeper' syntactic analyses also make it easier to link form and meaning, i.e. to derive semantics in addition to syntactic structure.

A research stay at CAS would therefore provide a unique opportunity to take stock of the available syntactico-semantic language resources, contrast their design, and use them as an empirical basis for a comparison between grammatical frameworks, in particular HPSG, LFG, but also of course the theoretically more advanced versions of dependency grammar (FGD and UD), as well as the graph-structured approach to meaning representation pursued in AMR. Open research questions in this context include:

(1.1) Which formal requirements on grammatical analysis can participants agree on, and which design aspects are mostly empirical questions that depend on language-specific or algorithmic needs?

(1.2) What devices exist in different formalisms, in particular HPSG, LFG, FGD, and UD, that deal with typological variation, and how do they relate to each other in the analyses of specific phenomena?

(1.3) To what extent it is possible to convert automatically between the different representations, and which conversions lose (by reduction) or add information, e.g. by 'injecting' linguistic theory?

(1.4) Which types and granularities of syntactico-semantic analysis can successfully contribute to different language technology applications today and to those of the foreseeable future?

5.2 Meaning Representation and Composition

Formal semantics, as developed in philosophy and theoretical linguistics, focuses on *meaning composition* in the sense introduced in Section 3.1 above. For this it draws on traditional tools from logic and set theory, following the seminal works of Montague (1974).

While the basic idea is clear and has remained constant since Montague, there are numerous open questions once we move from simple declarative clauses to more complex sentences, which often involve (hyper-)intensionality—i.e. attitudes, plans, beliefs, desires, etc.—and may involve reference to entities that do not actually exist. For example, our running example (1) appears to refer to (future) skiing events that do not exist at the moment of utterance. Intensionality raises many questions for theory, annotation, and application. On the theoretical side, it is unclear how to best model intensionality with the tools of logic. The traditional answer going back to Kripke (1963) and Hintikka (1969), still prevalent in theoretical linguistics, involves reference to *possible worlds*, but there are several alternatives on offer, some of which have been developed explicitly with computational applications in mind (e.g. Fox & Lappin, 2005). While this is a larger question that we will not address directly, it shows the wider philosophical and logical background for our research programme, and it also has clear implications for applications. For knowledge extraction tasks, for example, we may typically want to separate (parts of) utterances that make factual claims about the real world from ones that involve attitudes, plans, beliefs, etc. Such separation hinges on the notions of intensionality and the (only partially correlated) determination of veridicity (Karttunen & Zaenen, 2005).

Another complication is how to best capture obligatory referential dependencies of the kind we find between the three subject positions in (1), without being overly deterministic: It is well known that a verb like *promise*, although predominantly a subject-equi verb as illustrated in (1), can under certain circumstances ('controller shift') be an object-equi verb (2-a); in other contexts ('split control'), *both* arguments can be controllers (2-b).

(2) a. *Kim promised Sandy to be allowed to try skiing.*
  b. *Kim promised Sandy to go skiing together.*

It is a traditional assumption in LFG (Bresnan, 1982) that such structures involve anaphoric control (i.e. a pronoun-like element, whose reference is determined in ways similar to anaphora resolution; see Haug, 2013, 2014) while other control verbs involve structure sharing, i.e. 'reentrancy' in the sense of a single syntactic item filling several grammatical functions. The distinction is illustrated in the LFG functional structure in Figure 3, where the subject of *promise* and the subject of *try* are coindexed in the syntax, whereas the subjects of *try* and of its complement *ski* are structure-shared. Early work by Neidle (1982); Mohanan (1983); Andrews (1990) showed that this contrast accounts for a wide range of cross-linguistic distinctions in control behavior.

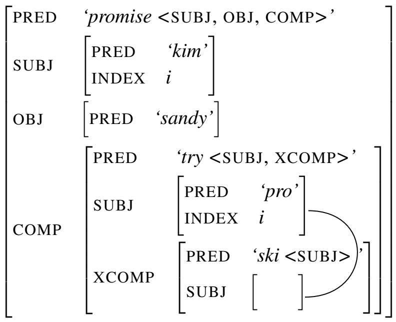

Figure 3: LFG f-structure for (1).

The distinction depends at least in part on the lexical semantics of the control predicate: in particular it has been argued that the *promise* but not *try* denotes a propositional attitude (Landau, 2013). This illustrates the more general fact that many of the most difficult questions in semantics crucially involve the interplay between the lexical semantics of the control predicate and the compositional semantics of the whole construction, a connection that that is hard to capture in the impoverished view of lexical meaning often employed in formal semantics. Recent work therefore tries to extend the type system traditionally used in compositional semantics so as achieve a more flexible composition system that can 'respond' to lexical meaning (see e.g. Asher, 2011; Luo, 2011).

A more radical approach to lexical semantics is the 'pure distributional' approach, based on collocations of content words represented as dimensionally reduced vectors, with linear algebraic operations like vector addition and multiplication substituting for traditional semantic composition in forming meanings for larger structures. Such representations are capable of representing the similarity of concepts as proximity in the multi-dimensional vector space, and hence potentially of detecting the similarity between paraphrases. However, it is hard to see how they can be interfaced with logical semantics, although this remains the subject of much research (Clark & Pulman, 2007; Copestake & Herbelot, 2013).

A second kind of distributional semantics seeks to identify relations of paraphrase and entailment directly in unseen text, using parsing or 'machine reading', and to build such logical relations into natural language semantics directly, treating paraphrases as clusters and entailment as logical conjunction (Lewis & Steedman, 2013). This approach has been shown to to be capable of capturing linguistically significant entailments, such as that *McCain regrets that he wasn't nominated* entails that *McCain wanted to be nominated*, which could be used to acquire the information that the semantics of verbs like *want* includes an implicit controlled subject of the complement *to be nominated*.

Another open issue is the nature of the syntax–semantics interface. Given the basic building blocks of word meanings and syntactic structures (like the ones in Figure 1 and Figure 3), a comprehensive syntactico-semantic theory must define the mechanisms that derive meaning representations, including adequate treatment of ambiguities in the mapping from syntax to semantics. These questions are actively being pursued in the HPSG and LFG communities, with foundational contributions by core members of our group (Dalrymple et al., 1993; Dalrymple, 1999; Copestake et al., 2001, 2005). A fundamental question for such approaches is how tight or loose the coupling between syntactic structure and semantic composition should be. Glue semantics—the dominant paradigm in LFG—assumes a categorial semantics with categories defined over labelled f-structures. Since the f-structure is a relatively abstract level of syntax (compare the tree in Figure 1 (top) with the structure in Figure 3), this leads to a relatively loose connection between semantics and surface syntax. On the other hand, categorial grammar—as the name implies—assumes a categorial structure in both syntax and semantics and hence a strict isomorphism between the two. Both approaches are represented in our group, giving us an opportunity to evaluate and contrast their predictions across phenomena and languages. Glue semantics can easily overgenerate ambiguities, while strict syntax–semantics isomorphism may be too strong a hypothesis for at least some phenomena, including in free word-order languages.

(2.1) What is the relationship between the grammatical properties of a certain construction and the lexical semantics of the items that take part in it? How can we adopt and extend known computational techniques to extract the relevant lexical information?

(2.2) How can distributional models be equipped with composition mechanisms for grammatical phenomena, such as control, with 'invisible' structure whose distributions are not directly observable?

(2.3) What are the trade-offs between loose and strict couplings of syntactic structure and semantic composition, and what constraints do these impose on different types of meaning representation?

6 Methodology and Expected Results

It is to some degree a defining property of the proposed research setup and its fluid residency at CAS in a somewhat distant future that our work programme cannot be organized just like a regular collaborative research project. The research questions sketched above are foundational and open-ended to a point that one cannot realistically expect conclusive answers to many of them in the foreseeable future. In fact, even if one were to assume the existence of an ultimate 'ground-truth' syntactico-semantic theory (with satisfactory computational properties), we would consider the zealous pursuit of a grand unification of contemporary perspectives an ill-advised short- to mid-term goal.

Instead, we view the range of related but distinct linguistic schools represented among group members as partly overlapping and partly complementary toolboxes. Dependency grammar, HPSG, and LFG each have been applied to diverse languages, have developed analyses for various linguistic phenomena, have been combined with different perspectives on semantics, and have each sought to somehow balance theory building and engineering. But to date, their respective inventories of languages and phenomena covered and design choices made at the syntax–semantics interface and in computational parsing diverge in both trivial and substantive ways—often bringing into focus different questions—but these remain (a) largely uninformed by each other and (b) inaccessible to researchers in applied language technologies.

To reap actual benefits from this diversity will require broad-minded but in-depth dialogue over an extended period of time. To obtain such cross-fertilization during our research stay at CAS, we will pursue a methodology of *phenomenon-oriented* and *data-driven* investigation. Concretely, we envision a handful of syntactico-semantic phenomena for the group to study collaboratively while at CAS, encompassing issues as foundational as argument vs. modifier identification and coordinate structures, for example, as well as somewhat more narrowly defined phenomena like comparatives, modal auxiliaries, nominalizations, and of course grammatical control. In the 'democratic' spirit of CAS, participants of our kick-off seminar (see below for details) will have the opportunity to jointly finalize this inventory.

To guide the exploration of these phenomena, we will curate selections of relevant exemplars for each phenomenon—in several languages—drawing on simplifications of corpus occurrences and on common linguistic examples. This work will result in a multi-lingual, partly aligned linguistic test suite organized by grammatical phenomena, which by itself will already provide a resource of longer-term value beyond our work at CAS. To instigate cross-framework comparison of candidate analyses for these exemplars, we will annotate relevant syntactic and semantic properties of the underlying phenomena, for example framework-specific assignments of argument and modifier types, or the gist of the logical relations involved in the analysis of the target constructions.

To relate this carefully curated, phenomenon-oriented test suite to naturally occuring language in context, and to connect the isolating theoretical study of constructed examplars with broad-coverage parsing using existing technologies, we envision two supporting activities. First, where applicable, target words (and phrases) will be paired with various distributional statistics obtained from very large text collections. Second, the group will also compile a small parallel corpus for at least a subset of the languages represented among participants. This corpus should draw on freely redistributable sources, for example existing translations of 'high-quality' text (for example tourism materials or technical documentation), but the group may decide to contract a limited amount of professional translation, giving specific instructions to translators on the overall goals of this effort. We expect that these activities will extend and inter-relate recent work within the LFG, HPSG, FGD, and UD communities on the creation of parallel treebanks, notably the ongoing ParGramBank and (Par)DeepBank initiatives (Flickinger, Zhang, & Kordoni, 2012; Flickinger, Kordoni, et al., 2012; Sulger et al., 2013).

The collaborative construction of these resources will in itself stimulate interaction and cohesion within the group, including the identification and definition of phenomena, selection of languages and exemplars, and their multi-framework annotation. And although we do not foresee constructing a very large resource—comprising maybe a set of a few hundred exemplars, at most—the multi-dimensional alignment of text and annotations will provide a solid foundation for contrastive investigation of the research challenges and questions identified in Section 5 above. In summary, the main discovery methods of the proposed group will be contrastive and experimental, aiming to (a) synthesize central insights from theoretical linguistics, to (b) evaluate their utility for automated parsing, and to (c) make this body of knowledge accessible to computer science researchers without much training in linguistics.

7 Mode of Operation and Dissemination

The start of our year at CAS will be marked by a kick-off seminar with all the researchers who are going to spend time in the group at some point during the year. The purpose of this workshop will be dual:

(a) to let everyone make their ongoing work and interests known to the rest of the group; and (b) to prioritize concrete research topics within the challenge areas described in Section 5 above, based on the specific questions raised there, but adapted to the state of the art in 2017. We expect each core group member to participate in work addressing both research challenges.

We anticipate that some of the group members will associate with a new Nordic collaboration in Web-Scale Natural Language Processing (NLPG), which was initiated in 2014 by, among others, Stephan Oepen and Joakim Nivre.[3] Through this association, group members with computational requirements that exceed CAS or university-level capacities will gain access to national high-performance computing facilities, including a specialized infrastructure and services for very large-scale experimentation with language data and for automated and standardized extrinsic evaluation of different parsing setups.

Work during the year at CAS will be organized through informal, 'open-space' meetings on specific research challenges, which will facilitate integration of the more junior post-doctoral and doctoral fellows participating from the inter-faculty SynSem initiative at the University of Oslo. Bi-weekly seminars for all present members of the group will be held to reflect on developments and results from these activities as they emerge. Where applicable, results will be documented in connection to the shared, on-line, and continuously growing catalogue of linguistic analyses for the phenomena targeted by the group.

We will organize two topical workshops during our period of residence at CAS: one smaller workshop by invitation towards the end of the first half year; and one high-profile international workshop towards completion of our stay at CAS. These outreach events will serve to ground our studies at CAS in related work and research communities, within Norway and abroad, and will further directly contribute to long-term international visibility and collaboration.

Finally, the ParGram and DELPH-IN initiatives have been holding (semi-)annual, one-week intensive working meetings since the mid-1990s, drawing theoretical linguists, grammar developers, and software engineers working on several different languages within these frameworks. The UD initiative is beginning to establish a similar meeting pattern in recent years. In the spirit of increased cross-fertilization between the linguistic theories represented in the group, we will propose that these networks (for the first time) hold a joint meeting in Oslo during our CAS residence. This meeting should be organized in conjunction with one of the two workshops, to reduce travel for overlapping participants and to also limit disruptions to our research time at CAS.

Among the expected outcomes of this configuration, we anticipate publications of theoretical, empirical, and practical results from our cross-cutting in-depth studies in top-tier international journals, as well as at least one jointly authored, high-visibility foundational reflection paper to be published in a major journal.

8 Budget Sketch

Cost Category	**Months**	**Cost**
Accomodation: Core Group	4+28	1,092
Accomodation: Visitors	8+8	696
Buy-Out and Substitutes	24	1,200
Kick-Off Seminar		100
Mid-Year Workshop		80
Dissemination Workshop		150
Running Expenses		120
Estimated Total		**3,438**

Table 1: Indicative budget (in kNOK).

Table 1 estimates the expenses for bringing together the proposed group for one year at CAS, starting in August 2017. Travel, accomodation, and other expenses related to temporary relocation to Oslo are budgeted using the rates of the Norwegian Research Council, at lump sums of kNOK 56 and 31 for the first and each following month, respectively. We expect that each member of the core group will on average spend eight months at CAS, while the duration of 'visits' by other international participants is projected to average two months. Although the exact schedule cannot be determined this far in advance, several participants plan on using sabbatical time for their stay at CAS; we estimate that about half of the four person years that international scholars spend with the group will require compensation to home institutions for teaching, administrative, and other duties. Exrapolating from the tentative indications by members of the core group, we budget kNOK 50 per person months for 'buying out' group members. Finally, we allocate between kNOK 80 and 150 for each of the three group-organized events (travel and accomodation for invited participants) and a discretionary kNOK 120 for general operating expenses of the group, e.g. publication or translation costs.

[3] See `http://www.mn.uio.no/ifi/om/aktuelt/arrangementer/andre/nlpg.html` for background.

References

Andrews, A. (1990). Case structure and control in modern icelandic. *Syntax and Semantics*, 187-234.

Asher, N. (2011). *The web of words*. Cambridge: Cambridge University Press.

Banarescu, L., Bonial, C., Cai, S., Georgescu, M., Griffitt, K., Hermjakob, U., … Schneider, N. (2013). Abstract meaning representation for sembanking. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse* (p. 178–186). Sofia, Bulgaria.

Baroni, M., Bernardi, R., & Zamparelli, R. (2014). Frege in space. A program for compositional distributional semantics. *Linguistic Issues in Language Technology*, 241–346.

Basile, V., Bos, J., Evang, K., & Venhuizen, N. (2012). Developing a large semantically annotated corpus. In *Proceedings of the 8th International Conference on Language Resources and Evaluation* (p. 3196–3200). Istanbul, Turkey.

Bender, E. M. (2011). On achieving and evaluating language-independence in NLP. *Linguistic Issues in Language Technology*(3), 1–26.

Bender, E. M., Flickinger, D., Oepen, S., Packard, W., & Copestake, A. (2015). Layers of interpretation. On grammar and compositionality. In *Proceedings of the 11th International Conference on Computational Semantics.* London, UK.

Bresnan, J. (1982). Control and complementation. *Linguistic Inquiry*, *13*, 343–434.

Bresnan, J. (2001). *Lexical-Functional Syntax*. Oxford: Blackwell.

Carpenter, B. (1992). *The Logic of Typed Feature Structures*. Cambridge, England: Cambridge University Press.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Clark, S., & Pulman, S. (2007). Combining symbolic and distributional models of meaning. In *Proceedings of the AAAI Spring Symposium on Quantum Interaction* (p. 52–55). Palo Alto, CA, USA.

Copestake, A., Flickinger, D., Pollard, C., & Sag, I. A. (2005). Minimal Recursion Semantics. An introduction. *Research on Language and Computation*, *3*(4), 281–332.

Copestake, A., & Herbelot, A. (2013). *Lexicalised compositionality.* Cambridge, UK. (Submitted for publication.)

Copestake, A., Lascarides, A., & Flickinger, D. (2001). An algebra for semantic construction in constraint-based grammars. In *Proceedings of the 39th Meeting of the Association for Computational Linguistics* (p. 140–147). Toulouse, France.

Dalrymple, M. (Ed.). (1999). *Semantics and Syntax in Lexical Functional Grammar: The resource logic approach.* MIT Press.

Dalrymple, M., Lamping, J., & Saraswat, V. (1993). *LFG semantics via constraints.* Utrecht, the Netherlands.

de Marneffe, M.-C., Dozat, T., Silveira, N., Haverinen, K., Ginter, F., Nivre, J., & Manning, C. D. (2014). Universal Stanford dependencies. A cross-linguistic typology. In *Proceedings of the 9th International Conference on Language Resources and Evaluation* (p. 4585–4592).

Flanigan, J., Thomson, S., Carbonell, J., Dyer, C., & Smith, N. A. (2014). A discriminative graph-based parser for the Abstract Meaning Representation. In *Proceedings of the 52nd Meeting of the Association for Computational Linguistics* (p. 1426–1436). Baltimore, Maryland, US.

Flickinger, D., Kordoni, V., Zhang, Y., Branco, A., Simov, K., Osenova, P., … Castro, S. (2012). ParDeepBank. Multiple parallel deep treebanking. In *Proceedings of the 11th International Workshop on Treebanks and Linguistic Theories* (p. 97–108). Lisbon, Portugal: Edições Colibri.

Flickinger, D., Zhang, Y., & Kordoni, V. (2012). DeepBank. A dynamically annotated treebank of the Wall Street Journal. In *Proceedings of the 11th International Workshop on Treebanks and Linguistic Theories* (p. 85–96). Lisbon, Portugal: Edições Colibri.

Fox, C., & Lappin, S. (2005). *Foundations of intensional semantics*. Oxford: Wiley-Blackwell.

Hajič, J., Hajičová, E., Panevová, J., Sgall, P., Bojar, O., Cinková, S., … Žabokrtský, Z. (2012). Announcing Prague Czech-English Dependency Treebank 2.0. In *Proceedings of the 8th International Conference on Language Resources and Evaluation* (p. 3153–3160). Istanbul, Turkey.

Haug, D. T. T. (2013). Partial control and anaphoric control in LFG. In M. Butt & T. H. King (Eds.), *Proceedings of LFG13* (p. 274–294). Debrecen, Hungary.

Haug, D. T. T. (2014). The anaphoric semantics of partial control. In M. W. Todd Snider Sarah D'Antonio (Ed.), *Semantics and linguistic theory* (Vol. 24, p. 213–233). LSA and CLC Publications.

Hintikka, J. (1969). Semantics for propositional attitudes. In J. W. Davis, D. J. Hockney, & W. K. Wilson (Eds.), *Philosophical logic* (pp. 21–45). Dordrecht: Reidel.

Hudson, R. A. (1984). *Word grammar*. Blackwell.

Ivanova, A., Oepen, S., Dridan, R., Flickinger, D., & Øvrelid, L. (2013). On different approaches to syntactic analysis into bi-lexical dependencies. An empirical comparison of direct, PCFG-based, and HPSG-based parsers. In *Proceedings of the 13th International Conference on Parsing Technologies* (p. 63–72). Nara, Japan.

Ivanova, A., Oepen, S., Øvrelid, L., & Flickinger, D. (2012). Who did what to whom? A contrastive study of syntacto-semantic dependencies. In *Proceedings of the Sixth Linguistic Annotation Workshop* (p. 2–11). Jeju, Republic of Korea.

Kaplan, R., & Bresnan, J. (1982). Lexical Functional Grammar. A formal system for grammatical representation. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (p. 173–281). Cambridge, MA, USA: MIT Press.

Karttunen, L., & Zaenen, A. (2005). Veridicity. In G. Katz, J. Pustejovsky, & F. Schilder (Eds.), *Annotating, extracting and reasoning about time and events.* Dagstuhl, Germany: Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany. Retrieved from `http://drops.dagstuhl.de/opus/volltexte/2005/314`

King, T. H., Crouch, R., Riezler, S., Dalrymple, M., & Kaplan, R. M. (2003). The PARC 700 Dependency Bank. In *Proceedings of the 4th International Workshop on Linguistically Interpreted Corpora* (p. 1–8). Budapest, Hungary.

Kripke, S. (1963). Semantical analysis of modal logic. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, *9*, 67–96.

Landau, I. (2013). *A two-tiered theory of control.* (to be published by MIT Press)

Lewis, M., & Steedman, M. (2013). Combined distributional and logical semantics. *Transactions of the Association for Computational Linguistics*, *1*, 179–192.

Luo, Z. (2011). Type-theoretical semantics with coercive subtyping. In N. Li & D. Lutz (Eds.), *Semantics and linguistic theory (SALT) 20* (pp. 38–56). eLanguage.

Mel'čuk, I. (1988). *Dependency syntax. Theory and practice*. Albany, NY, USA: SUNY Press.

Mohanan, K. P. (1983). Functional and anaphoric control. *Linguistic Inquiry*, *14*, 641–674.

Montague, R. (1974). *Formal philosophy. Selected papers of Richard Montague*. New Haven, CT, USA: Yale University Press.

Neidle, C. (1982). Case agreement in russian. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 391–426). Cambridge, MA: MIT Press.

Nivre, J. (2006). *Inductive dependency parsing*. Dordrecht, The Netherlands: Springer.

Nivre, J. (2015). Towards a universal grammar of natural language processing. In *Proceedings of the 16th International Conference on Intelligent Text Processing and Computational Linguistics.* Cairo, Egypt.

Oepen, S., Flickinger, D., Toutanova, K., & Manning, C. D. (2004). LinGO Redwoods. A rich and dynamic treebank for HPSG. *Research on Language and Computation*, *2*(4), 575–596.

Oepen, S., Kuhlmann, M., Miyao, Y., Zeman, D., Flickinger, D., Hajič, J., … Zhang, Y. (2014). SemEval 2014 Task 8. Broad-coverage semantic dependency parsing. In *Proceedings of the 8th International Workshop on Semantic Evaluation.* Dublin, Ireland.

Oepen, S., Kuhlmann, M., Miyao, Y., Zeman, D., Flickinger, D., Hajič, J., … Zhang, Y. (2015). SemEval 2015 Task 18. Broad-coverage semantic dependency parsing. In *Proceedings of the 9th International Workshop on Semantic Evaluation.* Denver, Colorado.

Plank, B., & van Noord, G. (2010). Grammar-driven versus data-driven. Which parsing system is more affected by domain shifts? In *Proceedings of the 2010 Workshop on NLP and Linguistics: Finding the common ground* (p. 25–33). Uppsala, Sweden: Association for Computational Linguistics.

Pollard, C., & Sag, I. A. (1987). *Information-based syntax and semantics. Volume 1: Fundamentals*. Stanford, USA: CSLI Publications.

Pollard, C., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. Chicago, USA: The University of Chicago Press.

Pustejovsky, J. (1991). The generative lexicon. *Computational Linguistics*(4), 409–441.

Sag, I. A., & Pollard, C. (1991). In integrated theory of complement control. *Language*(1), 63–113.

Sgall, P., Hajičová, E., & Panevová, J. (1986). *The meaning of the sentence and its semantic and pragmatic aspects*. Dordrecht, The Netherlands: D. Reidel Publishing Company.

Shieber, S. M. (1986). *An Introduction to Unification-Based Approaches to Grammar* (Vol. 4). CSLI Publications, Stanford, CA.

Sulger, S., Butt, M., King, T. H., Meurer, P., Laczkó, T., Rákosi, G., … Mistica, M. (2013). ParGramBank. The ParGram Parallel treebank. In *Proceedings of the 51th Meeting of the Association for Computational Linguistics* (p. 550–560). Sofia, Bulgaria.

Tesnière, L. (1959). *Éléments de syntaxe structurale*. Editions Klincksieck.

Xue, N., Bojar, O., Hajic, J., Palmer, M., Uresova, Z., & Zhang, X. (2014). Not an interlingua, but close. Comparison of English AMRs to Chinese and Czech. In *Proceedings of the 9th International Conference on Language Resources and Evaluation.* Reykjavik, Iceland.