

An HPSG-based Shared-Grammar for the Chinese Languages: ZHONG []

Zhenzhen Fan^{◇♣}

Sanghoun Song[♣]

Francis Bond[♣]

[◇]Institute of Systems Science, National University of Singapore

[♣]Division of Linguistics and Multilingual Studies, Nanyang Technological University
Singapore

zhenzhen@nus.edu.sg, {sanghoun, fcbond}@ntu.edu.sg

Abstract

This paper introduces our attempts to model the Chinese language using HPSG and MRS. Chinese refers to a family of various languages including Mandarin Chinese, Cantonese, Min, etc. These languages share a large amount of structure, though they may differ in orthography, lexicon, and syntax. To model these, we are building a family of grammars: ZHONG []. This grammar contains instantiations of various Chinese languages, sharing descriptions where possible. Currently we have prototype grammars for Cantonese and Mandarin in both simplified and traditional script, all based on a common core. The grammars also have facilities for robust parsing, sentence generation, and unknown word handling.

1 Introduction

Chinese is a group of related but sometimes mutually unintelligible languages that originated in China, including Mandarin Chinese, Cantonese, Min, etc. These languages have many grammatical similarities, though their orthography, vocabulary and syntax all differ from language to language. Thus, it is advantageous to implement a Chinese resource capable of covering both the common parts of the grammars and the linguistic diversity across the languages. Building an integrated grammar reduces the cost for resource construction and also helps the system reflect the genuine nature of the Chinese languages reliably.

This paper reports on our on-going project of building up an integrated computational grammar for these languages (ZHONG []) within the HPSG and MRS frameworks (Pollard and Sag, 1994; Copestake et al., 2005). The grammar is implemented using the collection of language process-

ing tools offered by the DELPH-IN (DEep Linguistic Processing with HPSG - INitiative, <http://www.delph-in.net>) consortium. This grammar combines a shared core for all the Chinese languages, as well as language specific descriptions. Currently we only have grammars for Mandarin Chinese (with simplified and traditional characters) and Cantonese, although we hope to add Min soon.

This paper describes how the grammar has been constructed and reports on its current capacity for parsing and generation. The paper is structured as follows: Section 2 offers background knowledge of the current work. Section 3 presents how the resource grammar works for the different Chinese languages. After discussing the specification of the grammar in Section 4, Section 5 conducts an evaluation to see coverage. Section 6 concludes this paper with an outlook for future work.

2 Background

2.1 Frameworks

The grammatical framework used for creating the Chinese shared-grammar is Head-driven Phrase Structure Grammar. HPSG models human language in a monostratal way via unification of constraints. Rules in HPSG are constructed as feature structures, which allows constructions to be analyzed via multiple inheritance hierarchies modelling the fact that constructions cluster into groups with a family resemblance that corresponds to a constraint on a common supertype.

The meaning representation system our grammar employs is Minimal Recursion Semantics. MRS representations have two significant characteristics. First, MRS introduces a flat representation expressing meanings by feature structures. Second, MRS takes advantage of underspecification for handling quantifier scopes and others, which allows flexibility in representation.

2.2 DELPH-IN

DELPH-IN is an informal collaboration between linguists and computer scientists adopting HPSG and MRS. DELPH-IN employs a shared format for grammatical representation based on type feature structures. The repository DELPH-IN readily provides consists of open-source tools, computational grammars, and language resources.

The tools include grammar development environment (LKB (Copestake, 2002)), efficient parsers/generators for language processing (PET (Callmeier, 2000), ACE (<http://sweaglesw.org/linguistics/ace>), *agree* (Slayden, 2012)), dynamic treebanking tools ([incr tsdb()]) (Oepen, 2001), ACE), machine translation engine (LOGON, ACE, *agree*), and stochastic models to select the most plausible interpretation. The collection of DELPH-IN grammars described in Type Definition Language include ERG for English (Flickinger, 2000), Jacy for Japanese (Siegel and Bender, 2002), GG for German (Crysmann, 2003), SRG for Spanish (Marimon, 2012), KRG for Korean (Kim et al., 2011), and others. The language resources contain test sets parsed with DELPH-IN grammars, such as the Redwoods Treebank in English (Oepen et al., 2004) and the Hinoki Treebank in Japanese (Bond et al., 2006), and a set of transfer rules (e.g. JaEn, (Bond et al., 2011)).

2.3 Previous Work on Chinese HPSG

2.3.1 Early Work

Early work on Chinese HPSG can be traced back to the 1990s, typically focusing on pure linguistic analysis of specific phenomena in Mandarin Chinese, such as the Chinese reflexive *ziji* (Xue et al., 1994), complement structure (Xue and McFetridge, 1996), and Chinese NPs (Gao, 1994; Xue and McFetridge, 1995; Ng, 1997).

Efforts towards a more comprehensive analysis of Mandarin Chinese in the framework of HPSG are documented in two PhD theses. The analysis in Gao (2000) covers topic sentences, valence alternations (including BA, ZAI, and other constructions), hierarchical argument structures, locative phrases, phrase structures, and resultative structures. The work of Li (2001) focuses more on the definition of word in Chinese for the problem of ambiguity in word segmentation, as well as two borderline problems between compounding/morphology and syntax - separable verbs and Chinese derivation and affixes.

2.3.2 Computational Grammars

In more recent work, in-depth analysis continues to be conducted on specific phenomena in Chinese HPSG, like the detailed account of Serial Verb Constructions (SVC) (Müller and Lipenkova, 2009), reanalysis of BA structure (Lipenkova, 2011), valence alternations and marking structures (Lipenkova, 2013), etc. However, the trend is to extend pure linguistic analysis to implementation of the grammar as a more general computational resource. This has led to a few independently developed HPSG grammars on Mandarin Chinese with MRS as the semantic representation format: ManGO (Yang, 2007), MCG (Zhang et al., 2011), and ChinGram (Müller and Lipenkova, 2013). ChinGram was implemented in the grammar development system TRALE (Meurers et al., 2002), whereas ManGO and MCG were developed using LKB and the LinGO Grammar Matrix customization system (Bender et al., 2010). These grammars cover a wide variety of core linguistic phenomena in Mandarin Chinese, but have limited lexical coverage as they typically only provide lexical entries for the words appearing in focused test suites. Yu et al. (2010), on the other hand, has explored a semi-automatic approach to developing a Chinese HPSG parser by proposing a skeleton design of the grammar and then learning a lexicon from an HPSG Treebank manually converted from the Penn Chinese Treebank 6.0 (Xue et al., 2005).

The foundation of our work is ManGO. Its test-suite is a Mandarin Chinese version of the MRS test suite used by the ERG, with short example sentences covering a wide range of phenomena such as intransitive, transitive, and ditransitive verbs, BA and BEI structures, clausal subjects/objects, aspect markers, prepositional and adverbial adjuncts, possessives, classifiers, numerals and determiners for noun phrases, predicative and attributive adjectives, locative and temporal phrases, nominalization, questions, imperative clauses, coordinations, etc. Its lexicon contains 231 lexical entries for 192 unique terms in 76 lexical types.

3 ZHONG []

The idea of letting different grammars share a common core to capture cross-linguistic generalization has been embraced by a number of projects as a more systematic approach for grammar development. The LinGO Grammar Matrix system (Bender et al., 2010) expedites the development of

complex grammars through grammar customization by providing a static core grammar that handles basic phrase types, semantic compositionality and general infrastructure. It also provides libraries for cross-linguistically variable phenomena, so that analyses of these can be dynamically generated as code based on user-configured parameters. The generated grammar is then extended usually manually by a grammar engineer. CoreGram (Müller, 2013) is motivated by a similar assumption that grammars sharing certain properties can be grouped into classes and thus share common files. Fokkens et al. (2012) proposes CLIMB (Comparative Libraries of Implementations with Matrix Basis), a methodology closely related to the LinGO Grammar Matrix. While still sharing implementation across different languages, the emphasis of CLIMB is facilitating the exploration and comparison of implementations of different analyses for the same phenomenon.

There’s also existing work sharing a common core grammar among languages within a language family. Avgustinova and Zhang (2009) builds a common Slavic core grammar (SlaviCore) shared by a closed set of languages in the Slavic language family. They further extended their work into SlaviCLIMB (Fokkens and Avgustinova, 2013), a dynamic grammar engineering component based on the CLIMB methodology, to capture language specific variations and facilitate grammar development for individual Slavic languages.

Extending the grammar development beyond Mandarin Chinese, ZHONG [] aims to provide a shared-grammar for Chinese and model various varieties of Chinese in a single hierarchy. The different Chinese grammars share some elements, such as basic word order, and separate other elements, such as lexemes and specific grammar rules (e.g., classifier constructions).

All grammars inherit from three common cores, viz. `zhong.tdl`, `zhong-lextypes.tdl`, and `zhong-letypes.tdl`. Building upon the common constraints. Mandarin and Cantonese inherit from `cmn.tdl` and `yue.tdl`, respectively. The distinctions between Mandarin and Cantonese captured so far include the expression of definiteness, classifiers, sentence final particles, aspect hierarchy, and some vocabulary. The Mandarin Chinese grammars are further divided into `zhs` and `zht` depending on whether the set of strings consists of simplified characters or tradi-

tional characters. These two further inherit from `zhs.tdl` and `zht.tdl`, respectively. The official webpage of ZHONG [], with demo and test results, is <http://moin.delph-in.net/ZhongTop>, and the entire data set can be freely downloaded from <https://github.com/delph-in/zhong>.

The size of the current grammar is presented in Table 1. ManGO, which ZHONG [] stems from, was created using the LinGO Grammar Matrix customization system. Hence, there are many fundamental types shared with the Grammar Matrix’s core (`matrix.tdl`).

Table 1: Size of grammar

| items | common | cmn | zhs | zht | yue |
|------------------------|--------|-----|--------|--------|-----|
| types | 383 | 21 | 1,017 | 7 | 17 |
| phrase rules | 79 | 0 | 0 | 0 | 0 |
| lexical rules | 69 | 4 | 3 | 0 | 2 |
| lex-entry types | 89 | 5 | 0 | 0 | 9 |
| lexicon | – | – | 43,067 | 17,470 | 903 |
| testsuites | – | – | 16 | 1 | 3 |

4 Components

4.1 Preprocessing and Postprocessing

ZHONG [] includes an unknown word handling module based on the chart-mapping technique of Adolphs et al. (2008). We have built a pipeline for converting raw text into a segmented POS-based lattice for input to the parser. The preprocessing stage for handling unknown words runs with the Stanford tools including the Chinese word segmenter (Tseng et al., 2005) and the Chinese Part-Of-Speech tagger (Toutanova et al., 2003). There are multiple different standards for segmenting the input string in Chinese, viz. Chinese Penn Treebank and Peking University. Between them, we are using the former because our fundamental development corpus NTU-MC (Tan and Bond, 2012) was segmented using that standard. We implemented a wrapper to run these tools in the pipeline using NLTK (Bird, 2006). In addition, the pre-processor includes some generic lexical entry rules for handling particular string patterns, such as numbers, dates, currency, emails, urls, etc. These lattice-based mapping rules work with a set of regular expressions. Building upon these two facilities, many lexical items not registered in the dictionary can be automatically identified and efficiently processed.

For postprocessing, we implemented a monolingual transfer grammar for paraphrasing simplified Mandarin Chinese, viz. `ZsZs`. This converts

MRS outputs in the parse results into more generic or more specific ones. Currently, this postprocessor works for generating intensifying constructions and classifier constructions.

4.2 Lexical Acquisition

As ManGO’s lexicon was small, our first task was to expand the lexical coverage of Zhong quickly. Our approach is to semi-automatically learn lexical entries from annotated corpora, starting from the sample of Sinica Treebank (**sinica**, Huang et al. (2000)) distributed with NLTK package and the Penn Chinese Treebank (**pctb**, Xue et al. (2005)) for Mandarin Chinese. Our main source at the beginning was **sinica** as it has a comprehensive set of POS tags, especially for verb subcategorization. Its POS tags were manually mapped to Zhong’s lexical types after careful study. Lexical entries for the mapped types were then created automatically. The tags from **pctb** are more coarse. We acquire words for the lexical types we are interested in by matching specific tree patterns against the treebank. The work is still ongoing.

As Zhong is used for both parsing and generation, we also try to learn additional information for the lexical entries, which is often required to constrain the grammar from generating unwanted sentences. For example, a list of classifiers (CL) can be readily learned from **sinica** and **pctb**. However, since in Mandarin Chinese there is a selective association between the sortal classifiers and the nouns, this association needs to be modeled so that during generation, a correct classifier can be selected for a certain noun. Our solution is to automatically build a frequency-based dictionary of noun-CL pairs, by extracting frequency information from a very large corpus. The corpus we used includes the latest dump of the Chinese Wikipedia, the second version of Chinese Gigaword (Graff et al., 2005), and the UM-Corpus (Tian et al., 2014). This data was cleaned, sentence delimited and converted to simplified Chinese script. It was further preprocessed using the Stanford Segmentor and POS tagger. Using very restrictive POS patterns, CL-noun pairs are extracted and filtered against a list of 204 sortal-CLs provided by Huang (Huang et al., 1997). They are then added into a lemma-based dictionary together with their frequency information. This lemma-based dictionary is further expanded into concept-based dictionary by mapping the lemmas to the concepts

in the Chinese Open Wordnet (Wang and Bond, 2013). The frequency information and possible CLs for matched senses are propagated to upper level through the union of CLs and respective sum of frequencies. Generation test on a set of held-out data reports a human validated performance of 88% on generation of classifiers using the concept-based dictionary and 80% using the lemma-based dictionary, whereas a baseline approach, taking \uparrow *ge* as the CL for every entry, gives 44.7%.

4.3 Configuration

ZHONG [] has been built up following the premise “parsing robustly and generating strictly” (Bond et al., 2008). This means that even a rather infelicitous sentence should be parsed, but the infelicitous sentence should be filtered out in generation. This different approach to parsing and generation can be facilitated using different configurations for compiling grammars. First, ZHONG [] includes a flag feature [STYLE *style*] for marking the felicity of particular lexical items and constructions, whose subtypes are *strict*, *robust*, *unproductive*, etc. Second, there are different types of roots: namely, `roots.tdl`, `roots-robust.tdl`, and `roots-strict.tdl`. The first one works for ordinary parsing and generation, the second one works with bridging rules to fill out the chasm between constructions, and the third one is particularly used for generation with the [STYLE *strict*] flag. Third, there are different scripts to load and compile the grammars within LKB and ACE, such as `config.tdl`, `config-robust.tdl`, and `config-strict.tdl`. The last one includes the list of items and rules that should be ignored in generation (`generation.ignore`).

For example, 去着 ‘go DUR’ may not sound good to Chinese native speakers, because the verb 去 tends not to co-occur with the durative aspect. Our grammar provides a parse tree for the sentence with a flag [STYLE *robust*] but does not generate such a sentence. To take another example, the punctuation markers are optionally treated in the ordinary and robust processing but obligatorily appear in the generation output produced by the grammar compiled by `config-strict.tdl`.

4.4 Grammar Enhancement

We have been enhancing the grammar with the objective to achieve coherent and consistent semantics constrained by syntax. Using the sentences

Table 2: Grammar enhancement

| | what we improved | what we added | what we plan to do |
|--------------------|--|---|--|
| grammar | topic-comment clefts BA and BEI NP structures classifiers argument structure adpositions | reduplication VV compounds A-not-A questions particles fragments interjections honorification | relative clauses nominalization serial verbs conjunctions |
| engineering | generation | unknown word handling bridging rules test modules full-forest treebanking | Wordnet incorporation transfer rules machine translation |

from the MRS testsuite, and supplemented by sentences collected from relevant literature and real corpus, we have improved the grammar on its handling of the known structures in the MRS testsuite, such as BA and BEI structures, NP structures, argument structures, classifiers, etc. At the same time, we have also created analyses to cover linguistically interesting phenomena new to the MRS testsuite, including reduplication of adjectives, resultative VV compounds, A-not-A questions, as well as the handling of particles, interjections, and fragments. Our work is summarized in Table 2.¹

4.5 Full-forest Treebanking

Using the simplified Mandarin Chinese grammar constructed thus far, we annotated two data sets by means of the full-forest treebanking tool (Packard, 2015). The data sets include the MRS Matrix testsuite in simplified Mandarin Chinese (<http://moin.delph-in.net/MatrixMrsTestSuite>) and the first 101 sentences in a novel (斑点带子案, *The Adventure of the Speckled Band* written by Arthur Conan Doyle, translated into Mandarin Chinese). The first set is a standard testsuite used in DELPH-In for testing grammars’ coverage of simple semantic phenomena. Of the 107 sentences, 102 can be parsed with the current grammar. Of these, 14 outputs were rejected in the annotation because no parse tree licenses the desired semantics. The second test suite was chosen because there exists a comparable annotated corpus written in four other languages (English, Spanish, Russian, and Korean) (Song, 2014). Because this is a running text consisting of longer sentences, the parse coverage is still poor: 12 out of 101. Of these 12, 8 were rejected for inadequate semantics. Annotating this running text, we learned that

¹The implementation of each grammatical phenomenon provided in Table 2 will be separately discussed in a series of upcoming papers.

the current grammar does not properly process relative clauses and serial verb constructions. These two phenomena are at the top of our agenda for grammar improvement.

5 Evaluation

We measured the coverage of the current grammar focusing on simplified Mandarin Chinese (abbreviated to zhs). We have two groups of test suites. First, we use three linguistic phenomena-based testsuites: the testsuite constructed at Free University of Berlin (**fu-berlin**, Müller and Lipenkova (2013)), the testsuite of the Mandarin Chinese Grammar (**mcg-wxl**, Zhang et al. (2011)), and the JEC basic sentences (**jec**, Kawahara and Kurohashi (2006)). Second, we use naturally occurring texts in order to check the computational feasibility of the current implementation. The corpora we used include the NTU-MC (**ntumc**, Tan and Bond (2012)), the Penn Chinese Treebank (**pctb**, Xue et al. (2005)), and the Sinica Treebank (**sinica**, Huang et al. (2000)). We used the entire NTU-MC (7,460 sentences) and extracted the first 5,000 sentences from the other two corpora. The tools for running tests are pyDelphin (<https://github.com/goodmami/pydelphin>) and gTest (<https://github.com/goodmami/gtest>). The result of coverage testing is provided in Table 3.

The numbers in parenthesis stand for the coverage of ungrammatical sentences. Note that only the first two include ungrammatical items. Since ungrammatical sentences had better be rejected, the smaller number means the better performance for those items. All the numbers in parenthesis are smaller than 5%, which shows that our grammar does not overgenerate very much.

When unknown word handling (**unk**) is facilitated, our current grammar provides relatively satisfactory results, as indicated in the third column.

Table 3: Coverage of simplified Mandarin (%)

| testsuite | plain | unk | br | unk+br | gen | end-to-end-success |
|------------------|--------------|--------------|--------------|--------------|-------|--------------------|
| fu-berlin | 22.22 (3.11) | 80.25 (3.12) | 22.22 (4.89) | 97.53 (4.97) | 90.91 | 20.20 |
| mcg-wxl | 57.28 (3.80) | 66.3 (3.78) | 82.44 (5.00) | 99.37 (5.00) | 92.94 | 53.24 |
| jec | 13.33 | 41.16 | 27.04 | 79.34 | 90.10 | 12.01 |
| ntumc | 3.47 | 15.58 | 10.54 | 47.82 | 70.54 | 2.45 |
| pctb | 0.84 | 7.10 | 10.18 | 43.70 | 42.86 | 0.36 |
| sinica | 3.88 | 40.36 | 6.52 | 65.00 | 80.41 | 3.12 |

However, the parsing coverage is still low when a running text is chosen for testing. Particularly, when it comes to the **pctb** testsuite, the coverage is only about 7%. There are two main reasons. First, the sentences in the **pctb** testsuite are much longer than those in the other testsuites. Second, our current grammar has not fully modeled relative clauses and serial verbs in Chinese, but the **pctb** testsuite includes many sentences containing such constructions. Thus, our immediate goal in grammar construction is to implement the constructions (see Table 2). When the **sinica** testsuite is used, the coverage is relatively high (40.36%). This is mainly because our lexical acquisition is mostly based on the corpus.

Using bridging rules (**br**) aims to facilitate robust parsing, which serves to minimize additional parsing costs (time and space) and maximize compatibility with existing platforms and tools. Since a set of bridging rules allows any two signs to combine into a phrase, the combination of unknown word handling and bridging rules (**unk+br**) provides the highest coverage, as indicated in the fifth column of Table 3. This implies that the **unk+br** mode enables our grammar to be used for training of statistical models and run-time applications in future work.

The generation coverage (**gen**) is calculated as follows: If a sentence is parsed, the MRS representation of the parse result is chosen as the input source for generation. Because the generation does not work with unknown word handling within the present infrastructure, the input source comes from the parse result of **plain**. If the generation process successfully produces one or more surface forms at the end, the generation coverage grows up. Notice that the generation coverage is not necessarily 100%, because the memory space for generation is limited (2GB in the current evaluation). The held-out testsuites result in more than 90% generation coverage, and the testsuites consisting of naturally occurring texts result in more than 70% except the **pctb** testsuite. We believe

that these measures are good for such a young grammar, although several challenging points remain. Finally, the end-to-end-success coverage from parsing to generation is measured by multiplying the values in the second column (**plain**) and the sixth column (**gen**).

6 Outlook

We will continue to enhance ZHONG [1] to handle the linguistic phenomena needed to parse our corpora (particularly, NTU-MC). Some of the tasks on the immediate agenda are: relative clauses, variations of nominalization, serial verb construction, conjunctions, other forms of verbal compounds, and more reduplication patterns. Lexical acquisition for `zht` and `yue` will also be performed to expand their lexical coverage.

We will also treebank other corpora, both as feedback to the grammarians and as a source of information on the distribution of phenomena (essential to training parse ranking models). As coverage increases we will exploit ZHONG [1] and other DELPH-IN grammars to build machine translation systems to and from Chinese.

Acknowledgments

We would like to express special thanks to Justin Chunlei Yang and Dan Flickinger for their enormous work on ManGO, which our current grammar is based on. In addition, we received much inspiration from Yi Zhang and Rui Wang and their Mandarin Chinese Grammar. We are grateful to Michael Wayne Goodman, Luis Mortado da Costa, Bo Chen, Joanna Sio Ut Seong, Shan Wang, František Kratochvíl, Huizhen Wang, Wenjie Wang, Giulia Bonansinga, David Moeljadi, Tuấn Anh Lê, Woodley Packard, Leslie Lee, and Jong-Bok Kim for their help and comments. Valuable comments from four anonymous reviewers are also much appreciated. Of course, we are solely responsible for all the remaining errors and infelicities. This research was supported in part by

the MOE Tier 2 grant *That's what you meant: a Rich Representation for Manipulation of Meaning* (MOE ARC41/13).

References

- Peter Adolphs, Stephan Oepen, Ulrich Callmeier, Berthold Crysmann, Dan Flickinger, and Bernd Kiefer. 2008. Some Fine Points of Hybrid Natural Language Parsing. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, and Daniel Tapias, editors, *Proceedings of the Sixth International Language Resources and Evaluation*, pages 1380–1387, Marrakech, Morocco.
- Tania Avgustinova and Yi Zhang. 2009. Parallel Grammar Engineering for Slavic Languages. In *Workshop on Grammar Engineering Across Frameworks at the ACL/IJCNLP*.
- Emily M. Bender, Scott Drellishak, Antske Fokkens, Laurie Poulson, and Safiyah Saleem. 2010. Grammar Customization. *Research on Language & Computation*, 8(1):23–72.
- Steven Bird. 2006. NLTK: the Natural Language Toolkit. In *Proceedings of the COLING/ACL on Interactive presentation sessions*, pages 69–72.
- Francis Bond, Sanae Fujita, and Takaaki Tanaka. 2006. The Hinoki Syntactic and Semantic Treebank of Japanese. *Language Resources and Evaluation*, 40(3–4):253–261.
- Francis Bond, Eric Nichols, Darren Scott Appling, and Michael Paul. 2008. Improving Statistical Machine Translation by Paraphrasing the Training Data. In *Proceedings of the International Workshop on Spoken Language Translation*, pages 150–157, Hawaii.
- Francis Bond, Stephan Oepen, Eric Nichols, Dan Flickinger, Erik Velldal, and Petter Haugereid. 2011. Deep Open-Source Machine Translation. *Machine Translation*, 25(2):87–105.
- Ulrich Callmeier. 2000. PET – a Platform for Experimentation with Efficient HPSG Processing Techniques. *Natural Language Engineering*, 6(1):99–107.
- Ann Copestake, Dan Flickinger, Carl Pollard, and Ivan A. Sag. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language & Computation*, 3(4):281–332.
- Ann Copestake. 2002. *Implementing Typed Feature Structure Grammars*. CSLI Publications, Stanford, CA.
- Berthold Crysmann. 2003. On the Efficient Implementation of German Verb Placement in HPSG. In *Proceedings of RANLP 2003*, pages 112–116, Borovets, Bulgaria.
- Dan Flickinger. 2000. On Building a More Efficient Grammar by Exploiting Types. *Natural Language Engineering*, 6(1):15–28.
- Antske Fokkens and Tania Avgustinova. 2013. Slav-iCLIMB: Combining Expertise for Slavic Grammar Development using a Metagrammar. In *Workshop on High-level Methodologies for Grammar Engineering*, pages 87–92.
- Antske Fokkens, Tania Avgustinova, and Yi Zhang. 2012. CLIMB Grammars: Three Projects using Metagrammar Engineering. In *Proceedings of the Eight International Conference on Language Resources and Evaluation*, pages 1672–1679, Istanbul.
- Qian Gao. 1994. Chinese NP Structure. *Linguistics*, 32:475–510.
- Qian Gao. 2000. *Argument Structure, HPSG, and Chinese grammar*. Ph.D. thesis, Ohio State University.
- David Graff, Ke Chen, Junbo Kong, and Kazuaki Maeda. 2005. *Chinese Gigaword Second Edition LDC2005T14*. Web Download. Linguistic Data Consortium.
- Chu-Ren Huang, Keh-Jiann Chen, and Ching-Hsiung Lai, editors. 1997. *Mandarin Daily Dictionary of Chinese Classifiers*. Mandarin Daily Press, Taipei.
- Chu-Ren Huang, Feng-Yi Chen, Keh-Jiann Chen, Zhao-ming Gao, and Kuang-Yu Chen. 2000. Sinica Treebank: Design Criteria, Annotation Guidelines, and On-line Interface. In *Proceedings of the Second Workshop on Chinese Language Processing*, pages 29–37, Hong Kong.
- Daisuke Kawahara and Sadao Kurohashi. 2006. Case Frame Compilation from the Web using High-Performance Computing. In *Proceedings of the 5th International Conference on Language Resources and Evaluation*, pages 1344–1347, Genoa.
- Jong-Bok Kim, Jaehyung Yang, Sanghoun Song, and Francis Bond. 2011. Deep Processing of Korean and the Development of the Korean Resource Grammar. *Linguistic Research*, 28(3):635–672.
- Wei Li. 2001. *The Morpho-Syntactic Interface in a Chinese Phrase Structure*. Ph.D. thesis, Simon Fraser University.
- Janna Lipenkova. 2011. Reanalysis of Obligatory Modifiers as Complements in the Chinese *ba*-Construction. In *Proceedings of the 18th International Head-driven Phrase Structure Grammar Conference, Stanford: CSLI Publications*.
- Janna Lipenkova. 2013. Valence Alternations and Marking Structures in a HPSG Grammar for Mandarin Chinese. In *Sixth International Joint Conference on Natural Language Processing*, pages 27–35.
- Montserrat Marimon. 2012. The Spanish DELPHIN Grammar. *Language Resources and Evaluation*, 47(2):371–397.

- W Detmar Meurers, Gerald Penn, and Frank Richter. 2002. A Web-Based Instructional Platform for Constraint-Based Grammar Formalisms and Parsing. In *Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics-Volume 1*, pages 19–26.
- Stefan Müller and Janna Lipenkova. 2009. Serial Verb Constructions in Chinese: An HPSG Account. In *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar*, pages 234–254.
- Stefan Müller and Janna Lipenkova. 2013. ChinGram: A TRALE Implementation of an HPSG Fragment of Mandarin Chinese. In *Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation (PACLIC 27)*, pages 240–249, Taipei, Taiwan. Department of English, National Chengchi University.
- Stefan Müller. 2013. The CoreGram Project: Theoretical Linguistics, Theory Development and Verification. *Ms. Freie Universität Berlin*.
- Say Kiat Ng. 1997. A Double-Specifier Account of Chinese NPs Using Head-Driven Phrase Structure Grammar. MSc Thesis, Department of Linguistics, University of Edinburgh.
- Stephan Oepen, Dan Flickinger, Kristina Toutanova, and Christopher D. Manning. 2004. LinGO Redwoods: A Rich and Dynamic Treebank for HPSG. *Research on Language & Computation*, 2(4):575–596.
- Stephan Oepen. 2001. [incr tsdb()] — Competence and Performance Laboratory. User Manual. Technical report, Computational Linguistics, Saarland University.
- Woodley Packard. 2015. Full Forest Treebanking. Master’s thesis, University of Washington.
- Carl Pollard and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. The University of Chicago Press, Chicago, IL.
- Melanie Siegel and Emily M. Bender. 2002. Efficient Deep Processing of Japanese. In *Proceedings of the 3rd Workshop on Asian Language Resources and International Standardization*, pages 1–8, Taipei.
- Glenn C. Slayden. 2012. Array TFS Storage for Unification Grammars. Master’s thesis, University of Washington.
- Sanghoun Song. 2014. *A Grammar Library for Information Structure*. Ph.D. thesis, University of Washington.
- Liling Tan and Francis Bond. 2012. Building and Annotating the Linguistically Diverse NTU-MC (NTU-Multilingual Corpus). *International Journal of Asian Language Processing*, 22(4):161–174.
- Liang Tian, Derek F. Wong, Lidia S. Chao, Paulo Quaresma, Francisco Oliveira, and Lu Yi. 2014. UM-Corpus: A Large English-Chinese Parallel Corpus for Statistical Machine Translation. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*. European Language Resources Association (ELRA).
- Kristina Toutanova, Dan Klein, Christopher D Manning, and Yoram Singer. 2003. Feature-rich Part-of-Speech Tagging with a Cyclic Dependency Network. In *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*, pages 173–180.
- Huihsin Tseng, Pichuan Chang, Galen Andrew, Daniel Jurafsky, and Christopher Manning. 2005. A Conditional Random Field Word Segmenter. In *Proceedings of the Fourth SIGHAN Workshop on Chinese Language Processing*, pages 169–171.
- Shan Wang and Francis Bond. 2013. Building the Chinese Open Wordnet (COW): Starting from Core Synsets. In *Proceedings of the 11th Workshop on Asian Language Resources, a Workshop at IJCNLP-2013*, pages 10–18, Nagoya.
- Ping Xue and Paul McFetridge. 1995. DP Structure, HPSG and the Chinese NP. In *Proceedings of the 14th Annual Conference of Canadian Linguistics Association*.
- Ping Xue and Paul McFetridge. 1996. Complement Structure in Chinese. In *Proceedings of the 32nd Annual Meeting of the Chicago Linguistic Society (CLS ’96)*, Chicago, IL.
- Ping Xue, Carl Pollard, and Ivan A Sag. 1994. A New Perspective on Chinese *ziji*. In *the Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*.
- Naiwen Xue, Fei Xia, Fu-Dong Chiou, and Marta Palmer. 2005. The Penn Chinese TreeBank: Phrase Structure Annotation of a Large Corpus. *Natural Language Engineering*, 11(02):207–238.
- Chunlei Yang. 2007. *Expert Systems for Pragmatic Interpretations of ziji and Quantified Noun Phrases in HPSG*. Ph.D. thesis, Shanghai International Studies University.
- Kun Yu, Yusuke Miyao, Xiangli Wang, Takuya Matsuzaki, and Junichi Tsujii. 2010. Semi-automatically Developing Chinese HPSG Grammar from the Penn Chinese Treebank for Deep Parsing. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 1417–1425.
- Yi Zhang, Rui Wang, and Yu Chen. 2011. Engineering a Deep HPSG for Mandarin Chinese. In *Proceedings of the 9th Workshop On Asian Language Resources*, Chiang Mai.