Syntactic Well-Formedness Diagnosis and Error-Based Coaching in Computer Assisted Language Learning using Machine Translation

Luis Morgado da Costa,[♠] **Francis Bond**[♠] **He Xiaoling**[◇]
[♠]Linguistics and Multilingual Studies [◇]Centre for Modern Languages
Nanyang Technological University, Singapore
<luis.passos.morgado@gmail.com,bond@ieee.org,xlhe@ntu.edu.sg>

Abstract

We present a novel approach to Computer Assisted Language Learning (CALL), using deep syntactic parsers and semantic based machine translation (MT) in diagnosing and providing explicit feedback on language learners' errors. We are currently developing a proof of concept system showing how semantic-based machine translation can, in conjunction with robust computational grammars, be used to interact with students, better understand their language errors, and help students correct their grammar through a series of useful feedback messages and guided language drills. Ultimately, we aim to prove the viability of a new integrated rule-based MT approach to disambiguate students' intended meaning in a CALL system. This is a necessary step to provide accurate coaching on how to correct ungrammatical input, and it will allow us to overcome a current bottleneck in the field — an exponential burst of ambiguity caused by ambiguous lexical items (Flickinger, 2010). From the users' interaction with the system, we will also produce a richly annotated Learner Corpus, annotated automatically with both syntactic and semantic information.

1 Introduction

Asserting if a sentence is grammatical or ungrammatical is, nowadays, a fairly easy task. The real challenge lies in answering questions like: *where is the error?*, *what is its correct form?*, or *what is its intended meaning?*. But especially in language learning environments (e.g. classrooms), where context is often poor, and students are requested to make up random sentences, context alone is usually not enough to answer the questions above. In fact, the pool of possible corrections of an ungrammatical sentence that arises from ambiguity (i.e. possible intended meanings) has been identified as the bottleneck of CALL systems (Flickinger, 2010), mainly because each possible meaning may trigger a different correction and explanation. Inside a traditional classroom, there is less of a problem since a human instructor can interact with the student to find the intended meaning. We propose to take advantage of high-quality machine translation (MT) and dialog-based computer-student interactions to similarly disambiguate learners' intended meaning and use this information to provide accurate and personalized grammar corrections and coaching. Consider the example below:

1. **That dog like the cat happy.*

The fact that (1) is ungrammatical is easy to determine. Existing computational grammars have been doing so for many years. The real challenge is answering questions like: *what is wrong with (1)?* or, *what is the correct form of (1)?*

Many systems struggle with these same questions. In a real life situation, context could possibly suffice to understand the utterance's intended meaning. But let us consider a situation where students were asked to make up a sentence that would make use of the words *dog* and *cat*. In this case, there is no context from which to extract clues about the intended meaning. It only seems natural that, should the teacher share another language with the student, they would make use of it to ask: *What did you mean?*

Proceedings of the 3rd Workshop on Natural Language Processing Techniques for Educational Applications, pages 107–116, Osaka, Japan, December 12 2016.

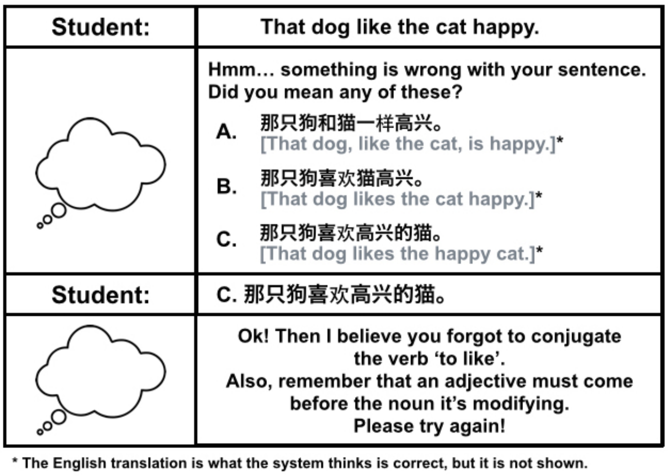


Figure 1: Semantic disambiguation example (Chinese speaker learning English)

We propose a new design of language tutoring systems that leverages on state-of-the-art NLP technology to provide explicit feedback on users' language errors. Figure 1 exemplifies the practical reach of the system we are building. The existence of many possible corrections of an ungrammatical sentence will trigger interaction between the system and the student. If there are multiple possible intended meanings, then it uses MT technology to ask what was meant, using the student's first language. After this, it can accurately provide hints about the errors. The use of MT in meaning disambiguation in CALL is, to the best of our knowledge, completely unprecedented, and it will enable systems to detect and provide feedback on many classes of grammatical errors with great confidence. In other words, our system leverages information across languages to find the exact intended meaning before correcting the student, helping push for a new state-of-the-art in CALL research. Venturing guesses, as is customary in these kinds of systems, can lead students into confusion, especially if the proposed correction had a different meaning than the one initially intended by the student.

The MT-Enabled Bilingual Online Language Tutor prototype we are developing focuses on entry level Mandarin as Second Language (L2) learners using English as their source language. At the end of the project, it will be evaluated, in a blended learning environment, by a large cohort of undergraduate students of the first level of Mandarin L2 at Nanyang Technological University. And while the goal of this project is to build a proof of concept system usable for early Mandarin learners, we are also catering for extensibility to further levels, bidirectionality, and even adding new languages in the future.

More technically, our system integrates precise syntactic parsers and semantics-based MT (Bond et al., 2005, 2011) to leverage information across languages. We are also integrating results from surveying word meanings and syntactic structures used by different levels of Mandarin L2, and the most common writing mistakes made by Mandarin L2 learners. This survey is guiding our design and implementation of mal-rules (‘error-production rules’), a type of grammar rule that selectively accepts ungrammatical sentences (Schneider and McCoy, 1998), but marking them as ungrammatical. These rules can then be used both to identify grammar errors and to reconstruct the semantics of ungrammatical inputs (Bender et al., 2004), which can then be used by the MT component to enable source-language interaction and feedback.

In the following section of this paper (2) we will discuss the motivation and significance of our system, followed by a survey of previous works in section (3). Section (4) will describe, in detail, the system implementation and our current implementation stage. Finally, we conclude and outline some future work.

2 New Learning Trends and Language Education

It is well known that Learning Sciences are rapidly entering a new era of online mediated education. A proof of this is the fact that many of the main players in the worldwide education system have identified the need of belonging to these new virtual learning spaces — joining existing or developing their own online learning platforms.

Concepts like Massive Open Online Courses (MOOCs) have only been around for a few years, and still this new learning paradigm has already caused an unprecedented change in worldwide education (Yuan et al., 2013). Unfortunately, the number of Massive Open Online Language Courses (MOOLC) available is proportionately very small — e.g. Perifanou and Economides (2014) report having found only 30 such courses in 2013. There is, easily arguable, not a lack of demand for such language courses, but a lack of technological infrastructure to support them.

And even though a considerable amount of research has been conducted in the last decades with regard to distance language learning, designing an efficient language learning course or developing a language learning platform is a very complex process. Perifanou and Economides (2014) states the ideas of **pedagogy** and **assessment** as being central challenges of these types of courses.

Our project was designed improve pedagogy — how the students learn a second language, based on improved feedback. It is targeted at university-level language learners, and will provide a scalable pedagogical infrastructure for online language learning. It can both be used in a blended learning environment, accompanying normal classroom style lectures, or it could eventually be further developed into a fully self-contained, self-paced online language course.

3 CALL: An overview

Artificial Intelligence (AI)'s contributions to CALL systems have, up to date, been mainly focused on problems like error classification, user modeling, expert systems, and Intelligent Tutoring Systems (Schulze, 2008; Gamper and Knapp, 2002). Following Gamper and Knapp (2002)'s survey on CALL systems, we known CALL systems differ mainly in the features they possess. Many of these systems have some domain knowledge, allowing detailed feedback to the learner, while others just guide students through a virtually designed course. Some present adaptive user models incorporating automated speech synthesis and recognition. Most systems use NLP techniques for analysis, but only a few also have generation capability. Some systems focus on one basic language skill (e.g. reading, writing, listening, or speaking), while others look for broader coverage. Some systems have a larger focus on grammar, others on vocabulary, and some even specialize in dialog interaction.

But ultimately, CALL systems are only a medium for language teaching. A study conducted by Nagata (1996) showed that it is not the medium itself (e.g. a computer, a book, etc.) that determines success in learning, it is the quality of the feedback produced by that medium that affects the results. This is why a language teacher is likely to be a better medium than a book, and the same reason why a properly designed CALL system can also be a better medium than a workbook, assuming that such systems can give valuable interactive feedback to the learner.

One of the main issues pointed out by Gamper and Knapp (2002) was that most CALL systems concentrate mainly on syntax and give less attention to semantic components, and only very few try to address the problem of pragmatics. At the same time, the integration of MT technology is also quite rare, and when is used, it mainly tries to give support to the training of translation skills.

More recently, a few CALL systems have started to use semantics to empower their precision and performance. An example of this is the adaptation of two high precision descriptive grammars (English and Norwegian) with semantic generation capacity into full-fledged CALL systems (Hellan et al., 2013; Flickinger and Yu, 2013; Flickinger, 2010; Bender et al., 2004). Both systems identify generation as crucial to their coaching feature. They apply the idea of semantically robust mal-rules, where the semantics of ungrammatical input is reconstructed by carefully designed rules that try to mimic common mistakes made by language learners.

The problem arises with the fact that each sentence allows a number of possible semantic representations (depending on the ambiguity of the lexicon and the strictness of the grammar rules). And while it is, in many cases, impossible to predict the intended meaning of the user's sentence, one solution is to make an educated guess with some statistical analysis (Hellan et al., 2013).

Still, previous systems (Hellan et al., 2013; Flickinger, 2010; Bender et al., 2004) report ambiguity as one of their central challenges. Balancing between the flexibility of the grammar and a high accuracy in disambiguating the intended analysis for each student sentence is essential to make the right diagnoses of errors (Flickinger, 2010). Even though statistical analysis may be tempting to solve this ambiguity, picking the incorrect intended meaning may mislead learners into thinking that they made an error that they did not. Up until today, the solution has been to reduce and control the lexicon in order to avoid ambiguity. The unfortunate consequence of this is that not all lexical entries can be equally represented in these systems.

4 Approach and Implementation

In this section we will describe the infrastructure and previous research grounding this project. We will also motivate and explain the concepts of graded lexical semantics and syntax, along with the choice of using the Open Multilingual Wordnet as a lexical ontology. We will finish by briefly introducing the relation between Learner Corpora and mal-rules, and their usage in CALL systems.

4.1 Integration of Descriptive Syntax and Semantics

The effort of creating a large-coverage, high-precision descriptive grammar is very time consuming (Copestake and Flickinger, 2000). For this reason, we chose to adapt existing grammars instead of creating new systems from scratch. Furthermore, we also needed the grammars to share some kind of semantic representation to be used for both parsing and generating across languages (i.e. to translate across languages using these grammars). We therefore selected grammars from the Deep Linguistic Processing with HPSG Initiative (DELPH-IN: Uszkoreit, 2002). DELPH-IN partners have agreed to and dedicated many years towards open-source multilingual parallel grammar development using Head Driven Phrase Structure Grammar theory (Sag et al., 2003) integrated with a computational semantics representation based on Minimal Recursion Semantics (MRS: Copestake et al., 2005; Copestake, 2007).

Concerning our specific languages of interest, for English we used the English Resource Grammar (ERG: Flickinger, 2000; Copestake and Flickinger, 2000), a grammar with a very large lexicon and wide coverage of syntactic phenomena; and for Mandarin Chinese ZHONG (Fan et al., 2015) [1] a more recent grammar with a solid coverage of core phenomena.

DELPH-IN grammars have been used in machine translation (Bond et al., 2005, 2011). DELPH-IN's MT work flow is based on semantic transfer systems — a source language is parsed by an HPSG grammar and a collection of underspecified semantic representations (i.e. MRSs) are generated and transferred to the target language's grammar — these then generate sentences encoding the same semantics in the target language.

4.2 Graded Lexical Semantics and Graded Parsers

Any CALL system must model some lexical semantics. Let us consider the example word *present*. Any mid or large sized dictionary of English would include multiple senses for this word. In the Princeton English Wordnet (Fellbaum, 1998), for example, there are 18 possible senses. Here is a selection of seven of those senses:

> 1. (noun) *something presented as a gift;* 2. (noun) *a verb tense that expresses actions or states at the time of speaking;* 3. (adjective) *being or existing in a specified place;* 4. (verb) *to give an exhibition of to an interested audience;* 5. (verb) *to introduce;* 6. (verb) *to give as a present; to make a gift of;* 7. (verb) *to recognise with a gesture prescribed by a military regulation;*

Considering the examples above, it is easy to acknowledge that lexical ambiguity is real. In real life situations, context is usually enough to disambiguate the intended sense. However, some of these senses are not commonly used in everyday situations, and can perhaps be ignored in the context of foreign language learning. Not even native speakers ever have full control of the lexical inventory of their language. So it is unreasonable to expect that an average user of English as a Foreign Language would be proficient using the word *present* as the verbal form *to recognise with a gesture prescribed by a military regulation*.

When considering many of the other common senses of *present*, it is important to note that language learners acquire different senses distributed in time (or language levels), either by necessity or by curricular requirement. Dewaele and Ip (2013) present a conclusive study that strongly relates Foreign Language Classroom Anxiety (FLCA) with Second Language Tolerance of Ambiguity (SLTA). Dealing with second language ambiguity is an important source of language use anxiety. This is also evident in the way most language courses are structured, as it is common practice to protect language learners from all the possible senses of a word until the language complexity so demands. Gradually learning to cope with

[1] http://moin.delph-in.net/ZhongTop

ambiguity is directly correlated with proficiency in any given language: the incremental aspect of this process is a very important notion to take into consideration.

Exploiting this incremental increase in ambiguity can be effectively used to minimize syntactic ambiguity. Descriptive grammars can be adapted to exploit this notion of natural gradual complexity of the learning process. Constraining the available lexicon by language levels, or even to the specific lexicon a user is known to possess, can help reduce ambiguity, by ignoring what the student could not have intended because it was out of their current knowledge.

We are building a model of language level based on a survey of the following resources: the first and second level of the Hanyu Shuiping Kaoshi (HSK) official language examinations, the first volume of the textbook New Practical Chinese Reader (Xun, 2010), Chinese Link: Beginning Chinese, Simplified Character Version, Level 1/Part 1 (Wu et al., 2010) and supplementary materials presented to the first level of Mandarin Chinese, as taught at our home institution.

These materials are being currently surveyed for their natural increment of lexical senses introduced to students. In addition, these same sources also contain information concerning syntactic complexity. In principle, the more complex syntactic structures are, the greater the likelihood that these would be introduced at later stages of language curricula. Following the same idea of gradually introducing lexical items into the descriptive grammars, we also argue that the same can be done with syntactic rules and constructions (e.g. minimize syntactic ambiguity by removing grammatical rules to which the student has not yet been introduced).

Both lexical and syntactic information is stored in a graded fashion in a database, relating statistical information about the syntactic structures' distribution across language levels, exams and curricula. This information will allow us to simplify descriptive grammars to a level of strictly necessary syntactic complexity to any surveyed language level — a system we named Graded Parser. By limiting the number of rules necessary to describe a specific language level, graded parsers can help avoid unnecessary ambiguity.

The surveyed lexical information will also be integrated in the Open Multilingual Wordnet (OMW) (Bond and Foster, 2013), a very large union of free wordnets. The OMW tightly integrates the English Princeton Wordnet (Fellbaum, 1998) and the Chinese Open Wordnet (Wang and Bond, 2013), allowing us to leverage on its structure to aid in the MT components.

4.3 Learner Corpora

First language transfer is widely accepted to play an important role in foreign language learning (Gass, 1988). Because of this, many CALL systems have been implemented for pairs of languages (i.e. a specific source language is considered in the development process) (Gamper and Knapp, 2002). CALL systems should be aware of the most common mistakes its users are known to make. For instance, missing the copula *to be* is a common mistake made by native Chinese speakers learning English (Schneider and McCoy, 1998). And, along the same lines, using an unnecessary copula (是 *shì*) in adjectival predication constructions is a common mistake made by native English speakers learning Mandarin.

The study of learner corpora focuses on the collection and analysis of language learner data. This data is especially of interest to CALL research if it has been error-tagged (i.e. all the errors in the corpus have been described with a set of tags: Granger, 2003). Before one can hope to design error detection and correction systems, it is necessary to survey errors contained in some learner corpora (Granger, 2003). Also, the appropriateness of the error correction in CALL systems is often measured against these kind of corpora (Schulze, 2008).

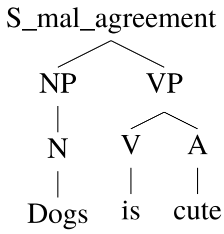
Even though producing an error-tagged corpus is very time-consuming, the huge return on invested resources is undeniable (Granger, 2003). For instance, documented and organized data can be used to customize the exercises in accordance with the learners' proficiency level and/or mother tongue background (Granger, 2003). Semantically annotated Learner Corpora are a good resource to predict the intended meaning of students (Hellan et al., 2013). And finally, the ungrammatical inputs collected by learner corpora can also be useful by providing examples of unparseable sentences for descriptive grammarians.

Many learner corpora are available for English learners coming from a Mandarin language background. Unfortunately, there seems to be an absence of readily available learner corpora made from Mandarin Chinese language learners. For this reason we are collecting and are currently annotating a learner corpus of Mandarin learners, using English as their source language. This learner corpus and the example sentences from the textbooks surveyed in Section 4.2 are being annotated using IMI (Bond et al., 2015), a multilingual semantic annotation environment that has been adapted to our needs.

4.4 Mal-rules

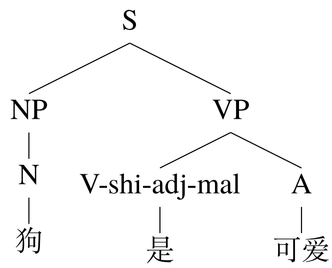
Mal-rules (also called 'error-production rules') (Schneider and McCoy, 1998) are a specific kind of rules that extend a descriptive grammar to make it accept (parse) ungrammatical phenomena. These mal-rules can be used to identify specific language errors, often triggering helpful messages to language learners. Consider the examples (2) and (3), below:

2. **Dogs is cute.*



3. *狗 是 可爱
   gǒu shì kě ài
   dog be cute

   "dogs are cute"



A descriptive grammar of English should reject (2) as a proper sentence. But if the intention were to capture the agreement error (between the subject NP and the VP), then expanding the English grammar with a mal-rule will serve this purpose perfectly. The node identified by *S_mal_agreement*, is a simple example of a mal-rule that was designed to explicitly allow a disagreement between the subject and the main clause. Similarly, a prescriptive grammar of Mandarin should reject (3) as a proper sentence, since the use of copula with adjectival predicates is not recommended (except in rare cases where pragmatics take a more prominent role). But as seen in (3), we can easily catch this error by adding a mal-lexical-entry, in this case named *V-shi-adj-mal* to flag the use of a copula that takes an adjectival complement. (3) is the first of many such mal-constructions to be implemented in our system, many more will follow.

The names of these rules are important, since checking the nodes of a parse tree can easily identify that the sentence was not grammatical because there is a *mal* (or any another tag) in the name of one of the nodes. The full rule name or lexical entry can be used to identify the specific kind of error and hence allow a system to say, for example, "there is something wrong with the agreement in that sentence", for (2), or "you should not use 是 before an adjective" for (3).

The mal-rules can be applied selectively. They can, for example, be used for parsing but not for generation (Bender et al., 2004), or to allow one type of error but not other. Also, because the grammars we are working with produce a semantic representation, these mal-rules are being designed to reconstruct the semantics of ungrammatical sentences, in a way that allows the generation of corrected counterparts (Bender et al., 2004). In some cases, the same error triggers multiple different mal-rules, each one reconstructing different semantics, so as to mimic different possible intended meanings by the student.

4.5 Our System's Architecture

In this section we will bring together all the previously presented details to elaborate on the design of our CALL system. Figure 2 presents a flowchart view of the coming description.

The final system will be web-based (accessible from any computer, tablet or phone with an internet connection). At the top of this system we have an authentication module. This ID will allow the system to retrieve all the necessary information to launch the tutoring system. The Student Model is the center of information. There we can find the vocabulary known by each student, the grammatical complexity the student is expected to work with, and the entire history of the student's interaction with the system (e.g. previous completed exercises, previous mistakes, time spent with each exercise, etc.).

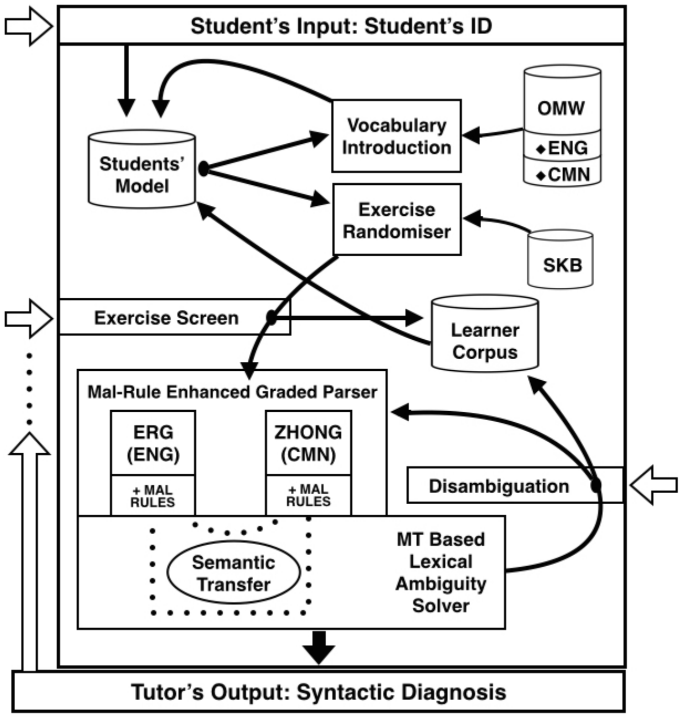


Figure 2: CALL system flowchart

Once the student is identified, the system allows two main tasks: Vocabulary Introduction and Exercise Randomiser. The Vocabulary Introduction module is directly linked to the new OMW extensions previously described (i.e. identifying individual lexical senses to specific language levels), allowing the student to preview necessary vocabulary in the target language. All the previously previewed vocabulary is stored in the Student Model and it feeds the Exercise Randomiser module.

The Exercise Randomiser module makes use of the previously known lexicon and the Syntactic Knowledge Base (SKB) to generate a one-sentence composition exercise, where learners must select words out of a randomly generated pool of words to compose a grammatical sentence. The number and type of spurious words that will be generated will be taken into account to determine its difficulty. The students inputted sentences will be stored in a Learner Corpus and sent to the Parser module. The Mal-Rule Enhanced Graded Parser module comprises the mal-rule enhanced grammars and a Semantic Transfer Machine Translation system.

The basic workflow of this module is as follows: if a parse is possible without activating mal-rules, then the solution is considered grammatical, the student is congratulated and the system returns to the beginning. If, on the other hand, one or more mal-rules are necessary to parse the student's solution, then there are two possible scenarios:

- there is no ambiguity about what errors were made: in this case the system can output a message prompting the student about the error made. The error tags will be added to the Learner Corpus, and the student will be asked to submit a new solution to the same problem until he/she can solve it; or
- there is ambiguity in the student's intended meaning, and different mal-rules that convey different meanings were triggered: in this case the system can't immediately output where the student made a mistake without first finding the intended meaning. In this case the solution of the student will enter the Ambiguity Solver module.

The MT Based Ambiguity Solver is a basic dialog system (similar to Figure 1) that will request help from the student to decide what the intended meaning was, and thus which errors were made. This module assumes that mal-rules have been constructed with robust semantics (i.e. reconstructing the right semantics for each particular error encoded in the rule). This will allow the system to use the reconstructed semantic representation to generate correct sentences. If there is a huge amount of ambiguity to solve, then parse ranking algorithms can help select the most probable set of intended meanings. This set of probable meanings is translated from target to source languages, and the student will be prompted to choose his/her intended meaning, between a set of translations.

Having found the student's intended meaning, a simple backwards analysis can be made to check which mal-rules were used to generate the selected choice. The ungrammatical solution will be stored in the learner corpus tagged with its intended meaning. The system will use the mal-rules used to generate the student's intended meaning in order to trigger an appropriate coaching message.

Finally, the system also has to account for completely unexpected sentences. Assuming the grammar will only be enhanced with common errors made by language learners, the system is likely to find inputs that it neither considers a grammatical input, nor does it have mal-rules that can help parse it. Also, students may still use a perfectly grammatical structure that the system is not expecting the student to use (given its graded architecture). In these cases, the system cannot say that the input is ungrammatical. Instead, a sentence that cannot be parsed can generate some general comments like: "You should not try to create sentences with structures you haven't learned yet. Try to make simpler sentences!". This is both uncommitted to the grammaticality of the input, and pedagogical in the sense that it tries to focus the student on his/her curriculum. These sentences should be flagged for the instructor to examine and give feedback on.

We hope to employ a few tricks to make the system friendlier to the students. For example, it can make intelligent use of ambiguous and unambiguous lexical entries to spare the student from having to take this disambiguation step too often. Also, when considering ambiguous input, it can automatically take the most probable intended meaning by default and output something like "If you mean A, then you need to be careful about mistake X. If you did not mean A, then help me understand what you meant by selecting from B, C or D."

The learner corpus compiled from this process will have very rich information when compared to other similar corpora. This system will not only collect statistics of common syntactic errors made by learners, but will also link these mistakes with the semantic annotation concerning the intended meaning. Also, when students are prompted to help the system disambiguate their solutions, an implicit parallel bilingual corpus is being created. The system will constantly be feeding itself information that makes it more intelligent, allowing interesting expansion over time.

5 Conclusions and Future Work

In this paper we have described a system that will hopefully help push for a new state-of-the-art on-line learning environments, by closely integrating semantics-based MT with computational grammars. Though still in an early stage of development, we have shown how we can use cutting-edge grammatical and semantic research research to build a system focused on reinforcing grammatical knowledge to Mandarin Chinese L2 learners.

Expandability (within the same language), adaptability to other languages and a component based architecture is at the core of our research agenda. So we expect not only that this system will be a useful resource for Mandarin L2 students, but that it can also help CALL research to further explore the integration of semantics, MT and other NLP field into its field.

At a pedagogical level, our approach empowers language educators, allowing them to focus their lectures on other major language skills (e.g. listening and speaking) rather than drilling. Educators can rely on CALL systems to provide personalized grammatical feedback to each individual student, and better attend to their individual struggles. At the same time, our system is designed for the students, providing them autonomy in self-paced study, and allowing them to spend more time practicing parts of the curriculum where they struggle the most.

We intend to evaluate this tool both with an intrinsic evaluation (how many errors can be correctly identified in a learner corpus) and an extrinsic one (in a classroom, does using this improve students test scores).

References

Bender, E. M., Flickinger, D., Oepen, S., Walsh, A., and Baldwin, T. (2004). Arboretum: Using a precision grammar for grammar checking in CALL. In *InSTIL/ICALL Symposium 2004*.

Bond, F. and Foster, R. (2013). Linking and extending an Open Multilingual Wordnet. In *51st Annual Meeting of the Association for Computational Linguistics: ACL-2013, Sofia*, pages 1352–1362.

Bond, F., Morgado da Costa, L., and Lê, T. A. (2015). IMI – a multilingual semantic annotation environment. *ACL-IJCNLP 2015*, pages 7–12.

Bond, F., Oepen, S., Nichols, E., Flickinger, D., Velldal, E., and Haugereid, P. (2011). Deep open-source machine translation. *Machine Translation*, 25(2):87–105.

Bond, F., Oepen, S., Siegel, M., Copestake, A., and Flickinger, D. (2005). Open source machine translation with DELPH-IN. In *Open-Source Machine Translation: Workshop at MT Summit X*, pages 15–22, Phuket.

Copestake, A. (2007). Semantic composition with (robust) minimal recursion semantics. In *Proceedings of the Workshop on Deep Linguistic Processing*, pages 73–80. Association for Computational Linguistics.

Copestake, A. and Flickinger, D. (2000). An open source grammar development environment and broad-coverage English grammar using HPSG. In *In proceedings of LREC 2000*, pages 591–600.

Copestake, A., Flickinger, D., Pollard, C., and Sag, I. A. (2005). Minimal recursion semantics: An introduction. *Research on Language and Computation*, 3(2-3):281–332.

Dewaele, J.-M. and Ip, T. S. (2013). The link between foreign language classroom anxiety, second language tolerance of ambiguity and self-rated English proficiency among Chinese learners. *Studies in Second Language Learning and Teaching*, (1):47–66.

Fan, Z., Song, S., and Bond, F. (2015). An hpsg-based shared-grammar for the chinese languages: Zhong [|]. In *Proceedings of the Grammar Engineering Across Frameworks (GEAF) Workshop*, pages 17–24.

Fellbaum, C. (1998). *Wordnet: An electronic lexical database*. MIT Press Cambridge.

Flickinger, D. (2000). On building a more effcient grammar by exploiting types. *Natural Language Engineering*, 6(01):15–28.

Flickinger, D. (2010). Prescription and explanation–using an HPSG implementation to teach writing skills. In *Invited talk, HPSG Conference*.

Flickinger, D. and Yu, J. (2013). Toward more precision in correction of grammatical errors. In *Proceedings of the Seventeenth Conference on Computational Natural Language Learning: Shared Task, CoNLL 2013, Sofia, Bulgaria, August 8-9, 2013*, pages 68–73.

Gamper, J. and Knapp, J. (2002). A review of intelligent CALL systems. *Computer Assisted Language Learning*, 15(4):329–342.

Gass, S. M. (1988). *Second Language Acquisition and Linguistic Theory: The Role of Language Transfer*, pages 384–403. Springer Netherlands, Dordrecht.

Granger, S. (2003). The International Corpus of Learner English: a new resource for foreign language learning and teaching and second language acquisition research. *Tesol Quarterly*, 37(3):538–546.

Hellan, L., Bruland, T., Aamot, E., and Sandoy, M. H. (2013). A Grammar Sparrer for Norwegian. In *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2013), Oslo, Norway. NEALT Proceedings Series*, volume 16.

Nagata, N. (1996). Computer vs. workbook instruction in second language acquisition. *CALICO journal*, 14(1):53–75.

Perifanou, M. and Economides, A. (2014). MOOCs for foreign language learning: an effort to explore and evaluate the first practices. *INTED2014 Proceedings*, pages 3561–3570.

Sag, I. A., Wasow, T., Bender, E. M., and Sag, I. A. (2003). *Syntactic theory: A formal introduction*, volume 2. CSLI Stanford.

Schneider, D. and McCoy, K. F. (1998). Recognizing syntactic errors in the writing of second language learners. In *Proceedings of the 17th International Conference on Computational Linguistics - Volume 2*, COLING '98, pages 1198–1204, Stroudsburg, PA, USA. Association for Computational Linguistics.

Schulze, M. (2008). AI in CALL – artificially inflated or almost imminent? *Calico Journal*, 25(3):510–527.

Uszkoreit, H. (2002). New chances for deep linguistic processing. pages XIV–XXVII, Taipei.

Wang, S. and Bond, F. (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.

Wu, S., Tian, W., and Zhang, Y. (2010). *Chinese Link: Beginning Chinese: Simplified Character Version, Level 1*. Chinese Link: Zhong Wen Tian Di. Beginning Chinese. Level 1. Prentice Hall.

Xun, L. (2010). *New Practical Chinese Reader Vol. 1 (2nd.Ed.)*. Beijing Language Culture University Press.

Yuan, L., Powell, S., and CETIS, J. (2013). MOOCs and open education: Implications for higher education. *Cetis White Paper*.