Grammatical error diagnosis in Fluid Construction Grammar

Beuls, Katrien

Published in:
Computer Assisted Language Learning

Publication date:
2014

Document Version
Peer reviewed version

This manuscript is the author's draft and has been published as:
Katrien Beuls (2014) Grammatical error diagnosis in Fluid Construction Grammar: A case study in L2 Spanish verb morphology, Computer Assisted Language Learning, 27:3, 246-260, DOI: 10.1080/09588221.2012.724426

Grammatical Error Diagnosis in Fluid Construction Grammar:

A Case Study in L2 Spanish Verb Morphology

Katrien Beuls

Vrije Universiteit Brussel, Brussels, Belgium

AI Lab, Pleinlaan 2, B-1050 Brussels

Katrien Beuls is currently pursuing a PhD in Computer Science from the University of Brussels (Belgium). She has a background in cognitive linguistics (University of Leuven, Belgium) and speech and language processing (University of Edinburgh, Scotland). Before joining the AI Lab in Brussels, Beuls worked as an intern at the Austrian Acoustic Research Institute (2008) and the Information Retrieval Facility in Vienna (2009), where she conducted research for her MSc dissertation on *Text categorization for Intellectual Property*. Beuls is particularly interested in building a prototype for tutoring games, in which a human user learns a language system from a software agent, inspired by the methodology of evolutionary language games. She received a four-year PhD grant from the Flemish Institute for Innovation and Technology (IWT).

Grammatical Error Diagnosis in Fluid Construction Grammar: A Case Study in L2 Spanish Verb Morphology

Construction grammar (CG) has been proposed as an adequate grammatical formalism for building intelligent language tutoring systems because it is highly compatible with the learning strategies observed in Second Language Learning. Unfortunately, the lack of computational CG implementations has made it impossible in the past to corroborate these proposals with actual language tutoring prototypes. However, recent advances in Fluid Construction Grammar (FCG) now offer exciting new ways of operationalizing robust and open-ended language processing within a construction grammar approach. This paper demonstrates its adequacy for CALL applications through a case study on error diagnosis in the domain of Spanish tense, aspect and modal morphology. The performance of the FCG tutor is tested on the Spanish Learner Language Oral Corpus (SPLOCC 2). This first FCG Spanish error diagnostic prototype achieves an accuracy of 70% on a total of 500 conjugation errors in four oral tasks carried out by 20 low intermediate and 20 advanced English learners of Spanish. Follow-up experiments will test this prototype on larger learner corpora of differing proficiency levels.

Keywords: Fluid Construction Grammar, Error analysis, Error diagnosis, Robust parsing, L2 Spanish verb morphology

Introduction

Construction grammar (CG) has been proposed as an adequate grammatical formalism for building intelligent language tutoring systems because it is highly compatible with the learning strategies observed in Second Language Learning. Unfortunately, the lack of computational CG implementations has made it impossible in the past to corroborate these proposals with actual language tutoring prototypes. However, recent advances in Fluid Construction Grammar (FCG) (Steels, 2011) now offer exciting new ways of operationalizing robust and open-ended language processing within a construction grammar approach.

FCG fulfils the two main tasks in a ICALL system that require NLP techniques, namely:

1. Rendering a structural interpretation of the learner's input even in presence of unexpected constructions (**robust parsing**); and
2. Identifying ungrammatical constructions and return targeted advice and an appropriate solution to the learner's error (**error diagnosis**) (Menzel & Schröder, 1999).

Construction Grammar (CG) has been proposed as an adequate grammatical framework that meets the three "criteria of adequacy" as proposed by Matthews (Matthews, 1993): **computational effectiveness** (successful computational implementation); **linguistic perspicuity** (descriptive power) and **acquisitional perspicuity** (contribution to explanation of L2 acquisition). Unfortunately, the lack of computational implementations has made it impossible to substantiate these proposals with operational models. Recent advances in FCG[1], however, break with these tendencies and offer new ways of building tutoring systems that model the cognitive operations employed in Second Language Learning (Gerasymova, Spranger, & Beuls, 2012). Schulze and Penner (2008) already discussed the potential of Construction Grammar (CG) "to overcome some challenges in ICALL and to facilitate a more thorough analysis of learner language in context and thus improve our knowledge about language learning processes" (p. 427).

This paper shows that FCG can rightfully be used as the language-processing component in an Intelligent CALL system. A case study on Spanish verb morphology (tense, aspect and mood) as it is produced by L2 native English learners of Spanish (low

[1] For recent developments and interactive webdemos the interested reader is referred to [URL left out for peer review].

intermediate level) in story telling tasks demonstrates the design and the performance of a possible FCG-based language tutor. Due to its double-layered architecture – that captures routine as well as meta-layer processing – FCG is a flexible formalism that is able to grasp the nature of learner errors and update its own knowledge (Beuls, van Trijp, & Wellens, 2012).

The first section of this paper homes in on typical approaches to error detection found in existing ICALL systems and the FCG-specific error detection and diagnosis. Section 2 discusses the FCG grammar that has been designed for the Spanish verb conjugation case study as well as the architecture that is needed for the error detection. This section is also accompanied by an interactive web demonstration ([URL left out for peer review]). Section 3 subsequently details the system's evaluation on a subcorpus of the Spanish Learner Language Oral Corpora (SPLLOC) (Mitchell, Dominguez, Maria, Myles, & Marsden, 2008) that contains hand coded verb form errors and the human tutor corrections. The results of the FCG conjugator can then easily be compared with the native human tutor corrections, which yields an accuracy of 58% (enhanced accuracy of 70%) on 500 tested verb form errors. Section 4 wraps up with conclusions and an outlook for future research.

2. Error Detection

The main source of information that informs a language tutor of the learner's knowledge level is the grammatical correctness of his or her interaction (or language game) with the tutor. Based on the categorization of the learner's input, a student model can be constructed. Misconceptions enter into student models "because learners are not domain experts and thus make errors" (Park Woolf, 2009: 56). Errors can thus be seen as benefits for the error analyses, or as Antos (1982) puts it: "errors reflect hypotheses of linguistic norms that learners form about the L2". In other words, errors can be seen as evidence for language learning processes and they are typically described as a "deviation from the norm" (Ellis, 1986: 51).

The norm to which the learner's input is compared is the constructions contained by the tutor's grammar, which in most cases equals the designer's grammar. An FCG tutor would receive a hand-written set of constructions to tackle the knowledge domain that is being tutored (e.g. German articles, English past tense, etc.) and expand this set over the course of the learning process as more learner input has been processed and potentially added constructions to deal with typical learner errors.

[Figure 1 near here]

Not every error is due to incorrect or inconsistent knowledge, but some errors are due to slips (Corder, 1967) such as fatigue, boredom, etc. Heift and Schulze (2007) also make a distinction between errors and errors. According to them, errors are a result of a lack of competence on behalf of the language learner. Errors, on the other hand, are performance-based "slips" such as typographical errors. The parser will detect both types, but the feedback that the learner receives should differ.

Typically, ICALL systems handle errors either by means of mal-rules (Weischedel, Voge, & James, 1978) or constraint relaxation (Murphy, Krüger, & Grieszl, 1988). In FCG, both methods can be intertwined since there is no predefined set of mal-rules that can catch typical learner errors in language processing but through the use of the meta-layer solutions are found to encountered processing problems and some kind of "mal-rules" can be learned by the system. Figure 1 includes a schematic representation of a possible FCG architecture that guarantees a robust parsing process of learner input.

3. FCG Architecture for Spanish Verb Conjugation

The FCG tutor architecture can be divided into three main blocks: classification of the verb form in the input (verb class, morpho-phonological pattern) (1), diagnosis of the learner error (2) and feedback on the type of error and the solution (3). In the classification phase, FCG analyses the verb form and adds the verb's paradigm to its inventory if it is not known yet. This makes it similar to the behavior of *The Spanish Verb* (Soria, 1997), that is however described as a "rather eclectic computational algorithm, rather than a fully-fledged linguistic formalism" (p. 43) and therefore rather the reverse of FCG.

Phases 2 and 3 concern the diagnosis and repair of learner errors, something that is captured by the FCG meta-level that monitors and repairs routine processing. The design of the meta-level operators that are used in tutoring requires a more elaborate implementation since feedback should not only include the corrected verb form but also the type of error that was committed.

Automatic Spanish verb conjugation

Spanish verb inflections are known to be highly complex: a single verb can potentially have 140 different verb forms when its full conjugational paradigm is taken into account: 19 tenses/moods, seven inflected forms per tense/mood, two infinitives, two gerunds and four participle forms (Bosque & Demonte, 1999). This high degree of complexity can easily lead to errors, especially for learners who are still acquiring the full dimensionality of the Spanish verb paradigm.

A Spanish verb form can generally be segmented into three parts: a verb stem, a tense-mood inflection and a person-number inflection. An example is the form "cant-aba-mos" (*we sang*) where "cant" is the stem from the verb "cantar" (*to sing*), "-aba" the tense-mood ending for the indicative past imperfect and "-mos" the person-number ending that indicates the subject is a second person singular entity. Figure 2 illustrates this segmentation visually.

[Figure 2 near here]

According to Rello & Basterrechea (2010), complexity in Spanish verb conjugation is due to four factors:

1. Unmarked person-number endings result in verb forms with only two overt parts available to the parser (e.g. "cant-aba"; *I/he/she sang*)";
2. There is a variety in stem realizations, i.e. the same verb can have more than one stem (e.g. "tener" => "ten", "teng", "tuv", "tendr");
3. Prefixes and suffixes can be added to the stem (e.g. "encantar", *to enchant*);
4. A verb can be irregular in that its stem, inflections or both deviate from the regular conjugation paradigm.

The automatic Spanish verb conjugator that is presented here is fully operational in FCG. It has been inspired by the commercially available *Onoma* system (www.onoma.es) (Basterrechea & Rello, 2010). The Onoma conjugator is implemented as a cascade of finite state transducers that implements a decision tree. The use of finite state transducers (FSTs) provides the possibility of generating verbal paradigms as well as the reverse process: the analysis of inflectional verb forms. Further, the use of a cascade structure facilitates the implementation of ordered alternation rules. *Onoma* can also handle neologisms in Spanish.

Departing from the verbal paradigms reported in (Basterrechea & Rello, 2010), a similar system has been implemented in FCG that creates new constructions for unknown infinitives based on a list of predefined patterns (including a range of exceptions). Unlike *Onoma*, our system is not coupled to a corpus for training but immediately departs from a fully specified decision tree. The FCG grammar is initialized with an inventory of baseline constructions, and is extended every time a new verb form is encountered based on the patterns. The baseline constructions are the following[2]:

1. Seven tense/aspect constructions to express the basic tense/aspect distinctions in Spanish: past perfect, past imperfect, past pluscuamperfect, present, present perfect, future, future perfect.
2. Two functional constructions that either create a main verb or an auxiliary verb out of a lexical item.
3. Four morphological constructions to express person-number endings (e.g. “-s”, “-mos”, etc.) and two unmarked morphological constructions for the unmarked endings in the 1st and 3rd person singular.
4. 24 morphological constructions to express tense-aspect-mood endings (e.g. “-aba”, “ieron”, “a”, “e”, etc.).
5. Two stress constructions that add a stress pattern to a finalized verb form in production. In general, the stress falls on the second but last syllable in Spanish words.
6. Nine phonological constructions that change instantiate the stress contour defined by the stress constructions (e.g. “-aba-“ => “ába”), add changes due to assimilation processes (e.g. “í” + “is” => “ís”), etc.
7. Six unmarked agreement constructions to handle the assignment of agent roles that is implicit in the verb forms (by default Spanish speakers do not express the subject overtly (by means of a pronoun)).

---

[2] The baseline constructions, as well as a limited number of added verbs are part of an interactive online demonstration that supports this paper. You find the demonstration on: [URL left out for peer review].

By means of illustration, one full parse of the verb form “encantaba” is included (*I/he/she was enchanted*). The infinitive “encantar” has been added to the baseline construction inventory. Before the form “encantaba” is passed to the FCG language processing engine, it is split based on the verb stems and endings present in the construction inventory into “encant” and “aba”. The FCG parse tree is included in Figure 3. A total of seven constructions apply when “encantaba” is parsed.

[Figure 3 near here]

First, the lexical construction matches on the verb stem (*encantar*). Then, the inflection triggers the application of the morphological construction for "aba" (*aba-past-imperfect-1*). The stem construction subsequently determines the stress pattern (regular) (*stem-accent-*1suffix) while the unmarked morphological construction defines that the person and number features are 1st or 3rd person singular (*1/3sg-morph*). The functional construction *verb-verbal* sets the function of “encantar” to a verbal use. The tense/aspect construction (*past-imperfect*) then creates a verb phrase that is characterized by a past imperfect meaning. Finally, the agreement construction (*agreement-3sg*) determines the verb form as having a 3rd person singular agent. An alternative solution would have been the 1st person singular agreement construction.

The resulting meaning looks as follows:

```
((3sg-agent ?agent ?context)
 (encantar ?event ?context)
 (time-sphere ?event ?time-point past)
 (event-overlaps ?time-point)
 (viewpoint ?event internal))
```

The last three predicates express the semantics of the tense/aspect suffix “-aba”: the event happened at a reference time point in the past and is reported from an internal perspective (imperfective; vs. external or perfective).

Error diagnosing

How are errors against verb morphology detected in the learner's input by the FCG tutor? There exist a number of errors learners can make when they conjugate verbs in Spanish. This section presents the FCG architecture that catches these errors, corrects them and returns the type of error that has been committed. The most commonly encountered errors are:

- **Verb class**: the inflectional ending belongs to a different verb class than the verb stem, e.g. *asustió (correction: asustó, *he/she scared*).
- **Tense, aspect, and mood**: the verb form is conjugated correctly but carries the wrong tense, aspect or mood, e.g. indicative instead of subjunctive.
- **Person, number** (agreement): the verb form is conjugated correctly but carries the wrong person or number feature, e.g. 1st person singular instead of 1st person plural.
- **Phonological**: the verb form is conjugated through regular patterns but should have undergone phonological changes such as assimilation, palatalization, fronting, etc., e.g. *juga (correction: juega, *he/she plays*).
- **Verb stem**: the verb stem is not part of the lexical inventory of the Spanish language due to errors in spelling, pronunciation or missing lexical knowledge, e.g. *descubir (correction: descubrir, *to discover*).

Of course, any combination of these errors can occur as well. For instance, a wrong verb form that is built on a non-existing verb stem and the wrong verb class such as in *descub-aba, where the correct stem would have been descubr- and the correct ending for the past imperfect 3sg form -ía, resulting in "descubría" (*he/she discovered*).

As has been outlined above, FCG is made up of a double-layered architecture to handle unexpected forms (parsing) or conceptualizations (production) in language processing. This architecture is included in Figure 4. This meta-layer is illustrated here for the phonological, verb class and verb stem error types. The goal is to diagnose errors and return this diagnosis together with the corrected verb form to the learner. Repair strategies currently produce the system's correction every time a particular learner problem is encountered. When we take this one step further, so-called "mal-rules" could be added in the repair stage. These rules could in the future immediately correct similar verb errors without having to run through the meta-layer again once a specific error has been diagnosed. A first step into this direction has been proposed by van Trijp (2012). The results of adding "mal-rules" to the Spanish verb conjugator falls outside the scope of this paper and will be discussed in future work.

[Figure 4 near here]

Detect learner/expert difference

This diagnostic checks whether at the end of a successful parse tree the expert would have said the same as the learner for the parsed meaning that the tree has rendered. An example is the parsing of "juga", which results in the meaning:

```
((3sg-agent ?agent ?context) (jugar ?event ?context)
(time-sphere ?event ?time-point present) (event-overlaps ?time-point)).
```

When this meaning is now re-produced by the expert system the utterance "juega" is returned. The diagnostic then checks the difference between the resulting linguistic feature structure after parsing "juga" and the result of producing "juega" and returns the non-matching features. The feature mismatch has occurred here in the phonological categorization of the verb stem. The vocalic nuclei differ when one compares the two verb stems: "u" vs. "ue". The vocalic nucleus mismatch is returned by the diagnostic.

Detect feature mismatch

This diagnostic detects a failure in the parse tree due to a feature mismatch. An illustrative example is the learner utterance “jugía”. Figure 5 contains the resulting FCG parse tree.

[Figure 5 near here]

After the verb stem has been parsed, the parsing of the inflected ending “ía” fails due to a failure in unification of both feature structures. The diagnostic detects which feature did not unify when the construction for `ía-past-imperfect-2/3` was merged with the linguistic structure that contained the verb stem of jugar: the verb class feature. The verb stem “jug” belongs to the 1st verb class, while the ending “ía” occurs only with stems of the 2nd or 3rd verb class. The verb class feature mismatch is returned by the diagnostic together with the corrected verb form (stem + ending). The ending could be retrieved from the FCG grammar since the diagnostic could tell use the correct verb class, tense, aspect and agreement features.

Detect unknown form

This diagnostic triggers when there is at least one string left unprocessed at the end of the parsing process. An appropriate example is here the utterance “juqaba” (correction: “jugaba”). The stem “juq” is unknown to the FCG expert system so parsing stops immediately since an inflection alone cannot lead to a successful parse. The diagnostic will then scan the FCG construction inventory for the verb stem that resembles the unknown form most closely. This is done based on a metric that relies on the Levenshtein distance with an additional weight for the first letter. The first letter is usually less error prone than the rest of a word when typical learner errors are examined (Yannakoudakis & Fawthrop, 1983; Sibley, Pollock, & Zamora, 1984).

The parsing process is then restarted with the verb stem that scored highest on the similarity test. The new form is now the correction “jugaba” and parsing will succeed. The diagnostic returns the corrected form and the fact that the verb stem the learner provided was unknown to the system. In a future implementation, the specific position in the stem that caused the failure (in this case “q” was used instead of “g”) could be returned as well in order to provide more precise feedback.

Remaining errors

Remaining errors are either agreement errors (different person or number) or the use of a different tense, aspect or mood than the one that was expected. Such errors can only be tracked if there is a target meaning available to the expert system. For instance, in exercises that have a well-defined solution. In such cases, a diagnostic to detect a meaning mismatch is appropriate. This FCG diagnostic compares the feature structure that results from parsing the learning utterance and the production result of the target meaning as soon as the parsed meaning differs from the target meaning. The mismatch in feature is then returned together with the corrected utterance. Combinations of multiple errors will automatically be captured since multiple diagnostics can trigger in a row and solve every error independently.

Example

Let us look at one example parse where the system classifies, diagnoses and repairs a verb form. The verb form *“pers**e**guió” (*he/she chased*) deviates from its correction (“pers**i**guió”) in one character: “e” > “i”. The first thing the FCG tutor does is analysing the input form. First of all, the infinitive is reconstructed based on a segmentation of the stem and its ending. It is checked whether the reconstructed infinitive already exists in the FCG grammar. Related forms such as “seguir” (*to follow*) can help to speed up the reconstruction process. If the original verb form is scrambled so much that no pattern can be detected, a closest match is taken from the existing infinitives. Future interactions can then restore potential early mistakes. Once the infinitive is found, the conjugation type can be deduced and the full verb paradigm constructed. According to the *Onoma* classification, “perseguir” belongs to group 4 (orthographical changes: “gu” > “g”) as well as group 5 (vocalic changes: “e” > “i”) (Basterrechea & Rello, 2010). The appropriate constructions that express these changes are added when they are still lacking in the FCG grammar.

The parsing process is subsequently started with the updated grammar. It does not fail until the final node of the search tree where the meta-level operators detect an instance of a *learner/expert mismatch* error. This could be detected by extracting the meaning of the final search node and producing this meaning with the expert grammar. A comparison of the two resulting feature structures yields a mismatch in the *vocalic nucleus* feature of the stem unit. The feedback can finally be generated based on this mismatch type and the expert production result (“persiguió”).

4. Evaluation

The Spanish Language Learner Oral Corpus (SPLLOC) was used to evaluate the error classification system. The SPLLOC is a second language learner corpus that contains exclusively oral data that has been collected and transcribed (CHILDES format) (MacWhinney, 1991) during the ESRC funded project "Linguistic development in L2 Spanish: creation and analysis of a learner corpus" (University of Southampton, University of Newcastle) (Mitchell et al., 2008). The material that has been collected includes learner narratives, interviews and picture description tasks. This material is freely available for use (www.splloc.soton.ac.uk). The corpus can be consulted in audio format (Wave, MP 3) and as transcribed text (transcription file, tagged file, XML).

The learner corpus

The SPLLOC 2 project is titled "The Emergence and Development of the Tense-Aspect System in L2 Spanish". The corpus contains samples of spoken Spanish produced by 60 instructed learners with English as their L1. The learners are all English L1 speakers who have learned L2 Spanish in educational contexts within the UK. Speakers from bilingual English/Spanish backgrounds or with extensive social contacts with Spanish speakers were excluded from the sample. According to the SPLOCC investigators, "it was not possible to control learner selection for gender as the large majority of L2 learners at college and university levels in the UK are female" (www.splloc.soton.ac.uk). Table 1 summarizes the levels present in the SPLOCC 2.

[Table 1 near here]

SPLOCC2 Tasks

The learners all undertook 5 tasks. "The first four tasks were designed to explore the learners' developing ability to describe past events orally in L2 Spanish in a variety of ways, and to relate these in sequences, in both more open and more controlled contexts (narrative tasks, guided interview, simultaneous events task)" (http://www.splloc.soton.ac.uk). All activities were undertaken individually with a member of the research team. The fifth task was a computer based interpretation task, and was specifically designed to explore learners' developing ability to distinguish different meanings of the Spanish imperfect and preterit. All task descriptions can be downloaded from the SPLOCC website.

Test Corpus for Error Analysis

For the evaluation of the FCG error analysis component we selected the lowest and the highest L2 learner level that was available in the SPLOCC 2 corpus: low intermediate (Lower secondary school) and advanced (University Year 4). For these two groups of 40 learners (20 per level), we used the four spoken tasks and extracted all errors automatically from the XML file, together with the native corrections. This results in 500 verb form errors. The majority of errors comes from the low intermediate group: 82% or 408 errors against 92 errors in the advanced learners subcorpus. This results in on average 20,4 errors/learner in the low intermediate group and 4,6 errors/learner in the advanced group.

Before the evaluation can start, the infinitive of each of the incorrect forms is added to the expert grammar so that the complete conjugation can be generated. The 500 verb forms can be traced back to a total of 93 infinitives. The FCG grammar contains 490 constructions to operationalize all conjugations. The most frequently used infinitives are: "perseguir" (36), "leer" (34), "jugar" (24), "haber" (22) "tranquilizar" (21) and "despertar" (21).

Evaluation Results

The FCG engine is given one verb form at a time that needs to be parsed. If it cannot be parsed with the constructions that are part of the grammar, the meta-layer catches the form and searches for a solution that can be parsed (either by transforming the input or by modifications to the transient feature structure, see above). The parsed meaning is then produced by the system and rendered to the learner together with a detailed report of the transformations that were made.

The performance of the FCG corrector is measured in terms of accuracy, which is defined as follows: Accuracy is the percentage of corrected forms that equals the human correction (given by the corpus). Table 2 gives an overview of the accuracy scores. The average score on the complete subcorpus (low intermediate and advanced) equals 58%. Group scores differ considerably: 52% of accurate corrections in the low intermediate group against 63% in the advanced group. Apart from the fact that the advanced errors only constituted 12% of the errors that were investigated, this also suggests that the types of errors made by more advanced learners are easier to detect by the algorithm.

[Table 2 near here]

The second row in Table 2 includes the results of the so-called "enhanced accuracy". Sometimes FCG corrections are accurate when you only consider the isolated form, without any notion of the discourse the form is embedded in. Table 3 gives some examples of such corrections paired with the original input form and the human correction. It occurs often that the person and number features are wrong (e.g. "habían" vs. "había") or that the input form does not contain an error and can be parsed accordingly (e.g. "contando"). These corrections are included in the "enhanced accuracy" count. With these corrections counted as accurate (due to a lack of information), the FCG corrector achieves a score of 70% enhanced accuracy. This time, the difference in accuracy has disappeared between the two levels.

[Table 3 near here]

In order to get an idea of the relevance of these results, the Microsoft Word spelling corrector was used to correct the list of verb forms, again without any embedding in a larger context. Figure 6 shows the results for all 500 errors, in terms of accuracy and enhanced accuracy. The difference in accuracy is almost 20% with FCG outperforming MS Word considerably.

[Figure 6 near here]

5. Conclusion

This paper summarized the results of a first demonstration of the suitability of Fluid Construction Grammar and its reflective architecture in the field of ICALL. The FCG framework supports robust parsing and offers the possibility to diagnose and repair learner errors together with offering constructive feedback. A case study on Spanish verb morphology has yielded initial results of 58% accuracy, which reaches up to 70% when corrections are included that are accurate given that the FCG engine does not have access to the context a verb form is encountered in.

Since the tested subcorpus was relatively small (500 errors, 94 verbs), future experiments should expand the list of learner errors. Moreover, when contextual information is taken into account, a semantic model could be constructed of the sentence or discourse structure the verb error is situated in. This would lead to better corrections for the agreement errors and verb forms that are grammatically correct but do not fit the discourse model (e.g. "contando" (*counting*) vs. "cantando" (*singing*)). FCG's potential applicability of mal-rules (van Trijp, 2012) that improve the meta-layer efficiency will also be explored in future work.

Acknowledgments

[Acknowledgements left out for peer review]

References

Antos, G. (1982). *Grundlagen einer Theorie des Formulierens. Textherstellung in geschriebener und gesprochener Sprache*. Tübingen: Niemeyer.

Basterrechea, E., & Rello, L. (2010). *El verbo en español. Construye tu propio verbo (The Spanish verb. Build your own verb)*. Madrid: Molino de Ideas.

Beuls, K., van Trijp, R., & Wellens, P. (2012). Diagnostics and Repairs in Fluid Construction Grammar. In L. Steels & M. Hild (Eds.), *Language Grounding in Robots*. New York: Springer.

Bosque, I., & Demonte, V. (1999). *Gramática descriptiva de la lengua española*. Espasa: Real Academia Española, Colección Nebrija y Bello.

Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, 159-170.

Ellis, R. (1986). *Understanding second language acquisition*. Oxford University Press.

Gerasymova, K., Spranger, M., & Beuls, K. (2012). A Language Strategy for Aspect: Encoding Aktionsarten through Morphology. In L. Steels (Ed.), *Experiments in Cultural Language Evolution*. Amsterdam: John Benjamins.

Heift, T., & Schulze, M. (2007). *Errors and Intelligence in Computer-Assisted Language Learning*. New York: Routledge.

MacWhinney, B. (1991). *The CHILDES project: tools for analyzing talk*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Matthews, C. (1993). Grammar Frameworks in Intelligent CALL. *CALICO Journal*, *11*(1), 5-27.

Menzel, W., & Schröder, I. (1999). Error diagnosis for language learning systems. *ReCALL*, (May), 20-30.

Mitchell, R., Dominguez, L., Maria, A., Myles, F., & Marsden, E. (2008). A new database for Spanish second language acquisition research. *EUROSLA Yearbook*, *8*(1), 287-304.

Murphy, M., Krüger, A., & Grieszl, A. (1988). RECALL - Providing an Individualized CALL Environment. In S. Jager, J. Nerbonne, & A. van Essen (Eds.), *Language Teaching and Language Technology* (pp. 62-73). Lisse: Swets and Zeitlinger.

Park Woolf, B. (2009). *Building Intelligent Interactive Tutors*. Morgan Kaufmann.

Rello, L., & Basterrechea, E. (2010). Onoma: a Conjugator Tool for Verbs and Verb Neologisms. *Procesamiento del Lenguaje Natural*, *45*, 129-137.

Schulze, M., & Penner, N. (2008). Construction grammar in ICALL. *Computer Assisted Language Learning*, *21*(5), 427–440. Taylor & Francis. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/09588220802447727

Sibley, E. H., Pollock, J. J., & Zamora, A. (1984). Automatic Spelling Correction in Scientific and Scholarly Text. *Communications of the ACM*, *27*(4).

Soria, J. (1997). Expert CALL:data-based versus knowledge-based interaction and feedback. *ReCALL*, *9*(2), 43-50.

Steels, L. (2011). *Design Patterns in Fluid Construction Grammar*. (L. Steels, Ed.). Amsterdam: John Benjamins.

Weischedel, R. M., Voge, W. M., & James, M. (1978). An Artificial Intelligence Approach to Language Instruction . *Artificial Intelligence*, *10*, 225-240.

Yannakoudakis, E. J., & Fawthrop, D. (1983). The rules of spelling errors. *Information Processing & Management*, *19*(2), 87-99.

van Trijp, R. (2012). A Reflective Architecture for Robust Language Processing and Learning. In L. Steels (Ed.), *Computational Issues in Fluid Construction Grammar*. Springer.